Rest-Tuning: Data-Efficient Adaptation of EEG Foundation Models to Individuals via Resting-State Signals

Yinuo Zhang[1,*] **Xinyu Fu**[1,4,*] **Qing Wang**[1] **Yinte Zhang**[1]
Junjie Yu[1] **Kexin Lou**[1,2,†] **Quanying Liu**[1,2,3,†]

[1] Department of Biomedical Engineering,
Southern University of Science and Technology, Shenzhen, China
[2] Omni-Intelligence, Shenzhen, China
[3] Shenzhen Loop Area Institute, Shenzhen, China
[4] Department of Biomedical Engineering,
South China University of Technology, Guangzhou, China

12531368@mail.sustech.edu.cn

[*] Equal contribution.
[†] Co-corresponding authors.

Abstract

Subject-specific distribution shifts make personalization a central challenge for EEG foundation models (FMs). We propose **Rest-Tuning**, a task-agnostic calibration framework that personalizes a population FM using just 1–3 minutes of unlabeled resting-state EEG. Rest-Tuning aligns masked student representations with intact-view targets from a frozen population teacher, combining contrastive discrimination with pointwise consistency to produce a reusable subject-calibrated backbone. Across two multi-subject datasets and four EEG FMs, Rest-Tuning outperforms task-only adaptation, achieving accuracy gains of up to 13.3 percentage points and reducing the required labeled-data fraction by over 40 percentage points on average. A single resting calibration also supports task adaptation across sessions up to two weeks later, extending its utility beyond the initial recording session. Calibration requires less than three minutes of computation per subject on average, and its cost is shared across subsequent tasks. Representation analyses show reduced subject-to-population distance, further LoRA interpolation analyses suggest improved rest–task and cross-task compatibility. These results establish **Rest-Tuning** as a **practical, fast, and reusable calibration interface** that translates population EEG FMs into personalized backbones for data-efficient decoding across tasks and sessions. Code is available at https://github.com/ncclab-sustech/Rest-Tuning.git

1 Introduction

Electroencephalography (EEG) provides non-invasive, portable, and temporally precise measurements of brain activity. Translating population-trained EEG models to individual users requires adapting to differences in anatomy, physiology, electrode coupling, and cognitive state [1–3]. These sources of variability alter the relationship between recorded signals and task labels, creating a distribution mismatch between a population model and a new subject (Figure 1a). Subject adaptation is therefore central to reliable EEG decoding [4–6].

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

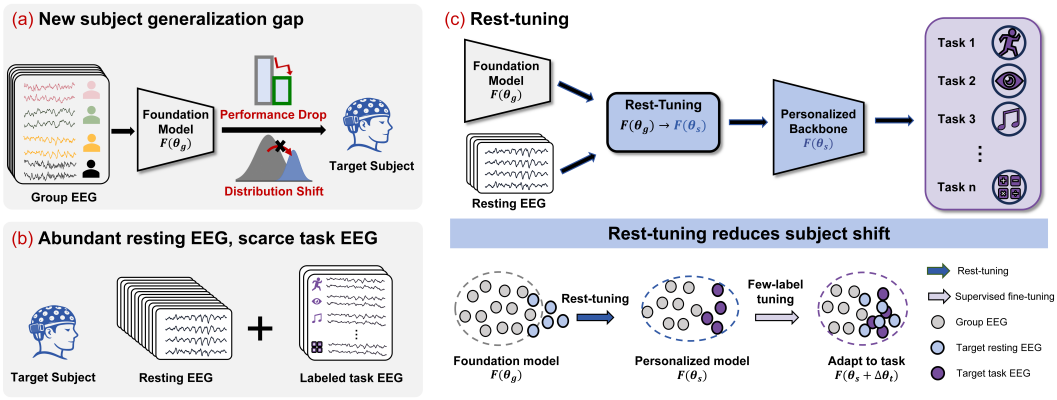

Figure 1: **Overview of Rest-Tuning.** (a) A population EEG FM F_{θ_g} encounters subject-specific distribution shifts when applied to a new user. (b) Labeled task EEG is scarce and task-specific, whereas resting-state EEG is comparatively easy to acquire and task-agnostic. (c) Rest-Tuning uses resting EEG to calibrate the backbone into F_{θ_s} before downstream supervision. The same subject-calibrated backbone can then be reused across multiple tasks with lightweight task-specific heads or adapters.

EEG foundation models (FMs) learn transferable representations from large collections of recordings [7–13]. Their population priors provide a useful starting point for downstream learning, while personalization typically uses labeled EEG from each subject–task pair. In this workflow, task supervision must support both adaptation to the individual and learning of the task decoder. Repeating the process across tasks increases acquisition and optimization costs (Figure 1b). A reusable subject representation offers a more efficient starting point: capture information about the individual once, then carry that calibration into subsequent task learning.

Resting-state EEG provides a natural signal for this purpose. It is easy to acquire without labels and contains subject-specific spectral, spatial, and connectivity structure associated with later task responses [14–18]. These properties make resting state suitable for a calibration stage that precedes task-specific supervision. Prior approaches, including BCM and ResTL, already demonstrate the value of baseline or resting EEG for transfer using source-subject rest and task recordings [19–21]. Population-pretrained FMs create a further opportunity: use the existing population representation as a reference and the target subject's rest as the sole calibration input, producing a backbone that can serve several downstream tasks.

We introduce **Rest-Tuning**, which separates subject calibration from task adaptation. A student initialized from the pretrained FM receives masked resting windows and learns from the intact-view representations of a frozen teacher. This design combines a fixed population reference with a learning signal drawn entirely from the target subject. The resulting checkpoint is retained as a common initialization for linear probing, task-specific LoRA, and full fine-tuning (Figure 1c). Each new task starts from the same calibrated backbone, making the initial resting calibration reusable.

Across two datasets and four EEG FMs, Rest-Tuning improves downstream decoding while reducing the labeled task data required for personalization. Effective calibration requires just 1–3 minutes of resting EEG, and a backbone calibrated once supports linear probing up to two weeks later on COGBCI. These findings establish resting EEG as a practical basis for reusable personalization: a short recording supplies a subject-specific initialization that benefits subsequent learning across tasks and sessions. Our contributions are:

- **Task-agnostic subject calibration.** Rest-Tuning uses unlabeled target-subject resting EEG to produce a personalized backbone reusable across downstream tasks.
- **Individualized rest–task alignment.** We show that Rest-Tuning reduces subject-to-population representation distances and improves rest–task and cross-task compatibility.
- **Data-efficient and reusable personalization.** Rest-Tuning matches task-only accuracy with an average reduction of over 40 pp in the labeled-data fraction and supports cross-session reuse of a

2 Related Work

EEG foundation models. EEG foundation models aim to learn reusable population representations that support diverse downstream decoding tasks. Models such as BIOT [11], LaBraM [10], EEGPT [12], and CBraMod [13] combine self-supervised pretraining on heterogeneous recordings with signal tokenization and spatial–temporal architectures tailored to EEG [22]. Their pretrained backbones provide a common starting point for downstream learning, separating the acquisition of shared EEG representations from adaptation to particular tasks and individuals. Translating these population priors into effective personalized models requires adaptation to subject-specific EEG distributions.

Masked and contrastive representation learning. Masked prediction and contrastive learning are widely used strategies for learning representations without task labels. In EEG, BENDR combines masking with contrastive prediction [23], MAEEG uses masked reconstruction [24], and EEG2Rep predicts masked content in latent space [25]. In vision, ConMIM formulates masked prediction as intra-image patch discrimination [26], while CMAE combines image reconstruction with instance-level contrastive learning [27], illustrating how masking and contrastive objectives can be integrated. Both generate representation targets from unmasked views using momentum-updated encoders. Rest-Tuning applies masked teacher–student learning to subject calibration on resting EEG, combining contrastive discrimination and matched-pair consistency against a frozen population-pretrained teacher.

Resting-state EEG for subject calibration. Prior work links resting-state connectivity to task-evoked activity, information transfer, and behavioral performance [14–17], motivating resting EEG as a source of subject-specific information for adaptation. BCM uses baseline EEG with task supervision to learn subject-invariant motor-imagery representations [19]. ResTL learns from source-subject resting and labeled task EEG, then adapts a motor-imagery classifier using target-subject rest [20]. Both approaches incorporate resting EEG into task-specific models.

Subject adaptation in EEG decoding. Subject adaptation is central to translating population-trained EEG models into effective individual-level decoders. Meta-learning learns adaptation strategies that transfer across subjects [28], while adapter-based methods reduce the number of trainable parameters required for personalization [5, 29]. For EEG foundation models, NeuroTTT addresses pretraining–downstream mismatch through domain-specific self-supervised fine-tuning and adapts to incoming EEG samples through test-time training [30]. Rest-Tuning complements these approaches with task-agnostic calibration on unlabeled target-subject rest, producing a personalized backbone reusable across downstream tasks.

3 Method

3.1 Problem setup

Our goal is to obtain a reusable subject-calibrated backbone for EEG FMs that can support multiple downstream tasks with limited labeled data.

Let F_{θ_g} be a pretrained EEG FM with population parameters θ_g. For a target subject s, we assume access to unlabeled resting-state EEG windows

$$\mathcal{D}^s_{\text{rest}} = \{x^r_i\}_{i=1}^{N_r}, \quad x^r_i \in \mathbb{R}^{C\times T}, \tag{1}$$

and limited labeled task data for one or more downstream tasks $\mathcal{T}$,

$$\mathcal{D}^{s,\mathcal{T}}_{\text{task}} = \{(x^t_j, y^t_j)\}_{j=1}^{N_t}, \quad N_t \ll N_{\text{population}}. \tag{2}$$

Rest-Tuning follows a two-stage personalization pipeline. First, an unlabeled subject-calibration stage maps the population parameters θ_g to subject-calibrated parameters θ_s using only $\mathcal{D}^s_{\text{rest}}$. Second, downstream task adaptation starts from F_{θ_s} and uses the labeled task data $\mathcal{D}^{s,\mathcal{T}}_{\text{task}}$. This separation is central to our evaluation: the task-only baseline uses the same downstream labels and adaptation interface, but starts from the original population model F_{θ_g} instead of the Rest-Tuned model F_{θ_s}.

3.2 Rest-Tuning

Teacher-student architecture. Rest-Tuning adapts the FM using a masked teacher-student objective on resting windows. Given a resting segment x_i^r, we construct a masked view

$$\tilde{x}_i^r = \mathcal{M}_\rho(x_i^r), \tag{3}$$

where $\mathcal{M}_\rho$ masks channels or temporal patches with ratio ρ. A trainable student, initialized from θ_g, encodes the masked view, while a frozen teacher keeps the original population parameters and encodes the unmasked view:

$$z_i^{(st)} = g\left(F_\theta(\tilde{x}_i^r)\right), \qquad z_i^{(te)} = \mathrm{sg}\left[g\left(F_{\theta_g}(x_i^r)\right)\right], \tag{4}$$

where $g(\cdot)$ is a parameter-free reshape operation and sg denotes stop-gradient. Only the student is optimized. The two branches have complementary roles: the student learns from an incomplete resting window, while the teacher supplies the population model's representation of the corresponding intact window. Asymmetric masking trains the student to predict the intact-view representation from partial observations. The frozen teacher keeps this prediction target tied to the population checkpoint throughout calibration on the subject's short recording. Appendix D illustrates the architecture.

Loss function. For a mini-batch of n resting windows ($n = 32$), we compute normalized embeddings $\hat{z}_i^{(st)}$ and $\hat{z}_j^{(te)}$ and form the student-teacher similarity matrix

$$S_{ij} = \hat{z}_i^{(st)\top} \hat{z}_j^{(te)}. \tag{5}$$

Each masked view is matched only to its own unmasked teacher view, giving an identity target distribution over the mini-batch. The contrastive loss is

$$\mathcal{L}_{\mathrm{rest}} = -\frac{1}{n} \sum_{i=1}^{n} \log \frac{\exp(S_{ii}/\tau)}{\sum_{j=1}^{n} \exp(S_{ij}/\tau)}. \tag{6}$$

Here, τ is the contrastive temperature. The masked contrastive form follows standard self-supervised representation learning, but its role here is different: it is applied as a per-subject, task-agnostic calibration stage rather than as population pretraining or downstream task adaptation [31–33]. To stabilize calibration and reduce overfitting to short rest recordings, we add a teacher–student consistency regularizer

$$\mathcal{L}_{\mathrm{mse}} = \frac{1}{n} \sum_{i=1}^{n} \left\| \hat{z}_i^{(st)} - \hat{z}_i^{(te)} \right\|_2^2, \tag{7}$$

with total loss

$$\mathcal{L}_{\mathrm{total}} = \mathcal{L}_{\mathrm{rest}} + \lambda_{\mathrm{mse}} \mathcal{L}_{\mathrm{mse}}. \tag{8}$$

We use $\lambda_{\mathrm{mse}} = 0.1$ as a fixed default. Optimizing this objective on $\mathcal{D}_{\mathrm{rest}}^s$ produces θ_s. Appendix C reports objective ablations and hyperparameter sensitivity.

3.3 Subject calibration and task adaptation interfaces

We perform subject-specific calibration using Rest-Tuning with two parameterizations that balance capacity and efficiency: (i) Rest-Full, which updates all backbone parameters and stores a full subject-specific checkpoint; (ii) Rest-LoRA, which inserts low-rank adapters into attention blocks while freezing the pretrained backbone, storing a compact subject-specific adapter. These Rest-Tuned models serve as the initialization F_{θ_s} for downstream task adaptation.

For downstream tasks, we evaluate three adaptation interfaces: (1) Linear probing, which freezes the calibrated backbone and trains a task head; (2) Task-LoRA, which trains task-specific low-rank adapters; (3) Task-Full, which performs full supervised fine-tuning. All task-only baselines employ the same task labels and interfaces but omit the Rest-Tuning stage. When Task-LoRA or Task-Full further modifies parameters for a given task, the original Rest-Tuned checkpoint is retained and reused as the initialization for future tasks from the same subject.

3.4 Subject-to-group alignment analysis

To test whether Rest-Tuning reduces the mismatch between an individual subject and the population representation, we perform a leave-one-subject-out alignment analysis. Since the original pretraining distribution is unavailable, we approximate the population prior within each dataset using the remaining subjects. For each held-out subject s_i, we first adapt the pretrained backbone on task EEG from the other $N-1$ subjects using the same self-supervised pretraining objective, excluding all data from s_i. This yields a subject-excluded population model $F_{\theta_{-i}}$.

In the baseline condition, both the held-out subject and the remaining subjects are encoded by $F_{\theta_{-i}}$:

$$P_{-i}^{\text{base}} = F_{\theta_{-i}}(X_{-i}), \qquad P_i^{\text{base}} = F_{\theta_{-i}}(X_i),$$

where X_i denotes EEG segments from the held-out subject and X_{-i} denotes EEG segments from the remaining subjects.

We then apply Rest-Tuning to $F_{\theta_{-i}}$ using only the held-out subject's unlabeled resting EEG, obtaining a subject-calibrated encoder $F_{\theta_i^{\text{rest}}}$. Both the held-out and population embeddings are re-encoded with this same encoder:

$$P_{-i}^{\text{rest}} = F_{\theta_i^{\text{rest}}}(X_{-i}), \qquad P_i^{\text{rest}} = F_{\theta_i^{\text{rest}}}(X_i).$$

The subject-to-population mismatch is quantified by the Wasserstein distance:

$$D_i^{\text{base}} = W(P_i^{\text{base}}, P_{-i}^{\text{base}}), \qquad D_i^{\text{rest}} = W(P_i^{\text{rest}}, P_{-i}^{\text{rest}}), \tag{9}$$

where $W(\cdot,\cdot)$ denotes the Wasserstein distance between embedding distributions. We repeat this procedure for every held-out subject. A decrease from D_i^{base} to D_i^{rest} indicates that Rest-Tuning moves the held-out subject closer to the population embedding geometry using only resting-state EEG.

3.5 Adapter-space interpretation

To analyze whether Rest-Tuning changes the parameter landscape for later task adaptation, we study linear interpolation in LoRA adapter space [34–36]. Let W_0 be the frozen backbone and ΔW_{rest} a Rest-LoRA update. For a downstream task, let ΔW_{task} be the task-supervised LoRA update. We evaluate the rest-to-task path

$$W(\alpha) = W_0 + (1-\alpha)\Delta W_{\text{task}} + \alpha \Delta W_{\text{rest}}, \quad \alpha \in [0,1]. \tag{10}$$

Here $\alpha = 0$ is the task-supervised endpoint and $\alpha = 1$ is the rest-calibrated endpoint. The random control replaces ΔW_{rest} with a randomly initialized adapter of the same rank. We also compare task-to-task paths between ΔW_{task_1} and ΔW_{task_2}:

$$W(\alpha) = W_0 + (1-\alpha)\Delta W_{\text{task}_1} + \alpha \Delta W_{\text{task}_2}. \tag{11}$$

These analyses test whether subject calibration makes rest- and task-derived adapters, or adapters from different tasks, more mutually compatible within the same subject.

4 Experiments

We evaluate Rest-Tuning through downstream decoding, representation analysis, cross-session reuse, and calibration efficiency. The decoding experiments measure the value of resting calibration under matched task supervision. Representation and adapter analyses characterize the calibrated model, while the cross-session experiment tests reuse over time. Ablations and efficiency analyses then establish the roles of the calibration objective, resting duration, and label budget. All task-only comparisons use the same downstream labels, splits, and adaptation interface.

4.1 Experimental setup

Datasets. M3CV [37] and COGBCI [38] provide resting and task recordings from the same subjects, supporting calibration before task learning. We use 95 labeled M3CV subjects from Session 1, covering motor, P300, steady-state sensory, and transient sensory tasks. The main COGBCI experiments use Session 1 from 29 subjects for NBack and MATB; the cross-session analysis additionally uses Sessions 2 and 3, one and two weeks later. Appendix A gives dataset and preprocessing details.

Table 1: **Uni- and multi-task** performance under 70% training data.

EEG Foundation Model	Rest-Tuning	Task-SFT	Trainable Params (Rest/Task)	M3CV				COGBCI			
				Unitask		Multitask		Unitask		Multitask	
				$\overline{\text{ACC}}$(%)	ΔACC(%)	$\overline{\text{ACC}}$(%)	ΔACC(%)	$\overline{\text{ACC}}$(%)	ΔACC(%)	$\overline{\text{ACC}}$(%)	ΔACC(%)
LaBraM	/	Task-LoRA	0 / 5.00%	56.96	/	36.04	/	60.46	/	56.73	/
	Rest-LoRA	Task-LoRA	5% / 5%	64.34	↑7.38	41.94	↑5.90	63.53	↑3.07	61.20	↑4.47
	Rest-Full	Task-LoRA	100% / 5%	64.74	↑7.78	42.90	↑6.86	68.94	↑8.48	66.71	↑9.98
	/	Task-Full	0 / 100%	61.66	/	45.95	/	72.58	/	73.44	/
	Rest-Full	Task-Full	100% / 100%	68.36	↑6.70	49.79	↑3.84	75.52	↑2.94	78.13	↑4.69
BIOT	/	Task-LoRA	0 / 4.5%	65.56	/	50.04	/	90.14	/	88.94	/
	Rest-LoRA	Task-LoRA	4.5% / 4.5%	70.57	↑5.01	52.24	↑2.20	90.98	↑0.84	89.50	↑0.56
	Rest-Full	Task-LoRA	100% / 4.5%	68.02	↑2.46	50.59	↑0.55	91.21	↑1.07	89.58	↑0.64
	/	Task-Full	0 / 100%	67.46	/	50.74	/	90.39	/	89.38	/
	Rest-Full	Task-Full	100% / 100%	69.27	↑1.81	52.82	↑2.08	91.24	↑0.85	90.45	↑1.07
CBraMod	/	Task-LoRA	0 / 3.03%	48.02	/	24.18	/	47.76	/	38.63	/
	Rest-LoRA	Task-LoRA	3.03% / 3.03%	50.46	↑2.44	24.95	↑0.77	60.95	↑13.19	40.24	↑1.61
	Rest-Full	Task-LoRA	100% / 3.03%	55.45	↑7.43	27.35	↑3.17	61.01	↑13.25	50.15	↑11.52
	/	Task-Full	0 / 100%	58.89	/	34.36	/	71.80	/	70.59	/
	Rest-Full	Task-Full	100% / 100%	62.49	↑3.60	37.55	↑3.19	75.40	↑3.60	73.69	↑3.10
EEGPT	/	Task-LoRA	0 / 3.03%	52.03	/	27.62	/	54.77	/	48.59	/
	Rest-LoRA	Task-LoRA	2.03% / 2.03%	54.54	↑2.51	27.96	↑0.34	54.92	↑0.15	49.17	↑0.58
	Rest-Full	Task-LoRA	100% / 2.03%	54.88	↑2.85	28.55	↑0.93	54.99	↑0.22	50.19	↑1.60
	/	Task-Full	0 / 100%	54.14	/	26.81	/	55.05	/	48.14	/
	Rest-Full	Task-Full	100% / 100%	55.64	↑1.50	30.30	↑3.49	57.47	↑2.42	54.69	↑6.55

Backbones and adaptation. We evaluate LaBraM [10], BIOT [11], CBraMod [13], and EEGPT [12]. Rest-Full updates the backbone with mask ratio $\rho = 0.3$; Rest-LoRA uses rank-8 attention adapters and $\rho = 0.7$. These are calibration-specific defaults evaluated in Appendix C.3 For downstream evaluation, we compare task-only adaptation against Rest-Tuning followed by the same task interface. Unless otherwise stated, results are averaged across target subjects and reported as mean $\pm$ standard deviation.

Protocols. The main evaluation uses subject-specific, within-session training, validation, and test partitions fixed with seed 42. Paired methods use identical task samples and splits. Resting EEG is used only for calibration and supplies no task labels. In *uni-task* decoding, each task is decoded separately in its original label space and accuracies are averaged across tasks. In *multi-task* decoding, one classifier predicts the disjoint union of task-specific labels constructed by the mapping in Appendix A. The main label budgets are 10% and 70%. Accuracies are reported as mean $\pm$ standard deviation across subjects and the cross-session is described in Section 4.4.

4.2 Rest-Tuning improves decoding across tasks and FMs

We first evaluate whether resting calibration improves downstream decoding while keeping the task interface fixed. Figure 2 reports linear probing accuracy across six downstream tasks and four EEG FMs. Rest-Tuning consistently outperforms the no-rest baseline, indicating that it enhances the backbone representations and improves task separability even before parameter-heavy supervised adaptation.

Table 1 summarizes uni-task and multi-task results for each backbone. Both Rest-LoRA and Rest-Full consistently improve task-only adaptation performance across tasks and backbones, with gains reaching approximately 10% in multiple tasks. These improvements persist across different task interfaces, highlighting that Rest-Tuning serves as a general subject-calibration stage rather than a task-specific adjustment.

All results above are obtained using 70% of the task labels for training. We further evaluate extreme low-label scenarios (train 10%), showing that backbone optimization (linear probing) and task adaptation gains remain robust (see Figure A7 and Table A15). Appendix N reports detailed task-level accuracy and F1 scores.

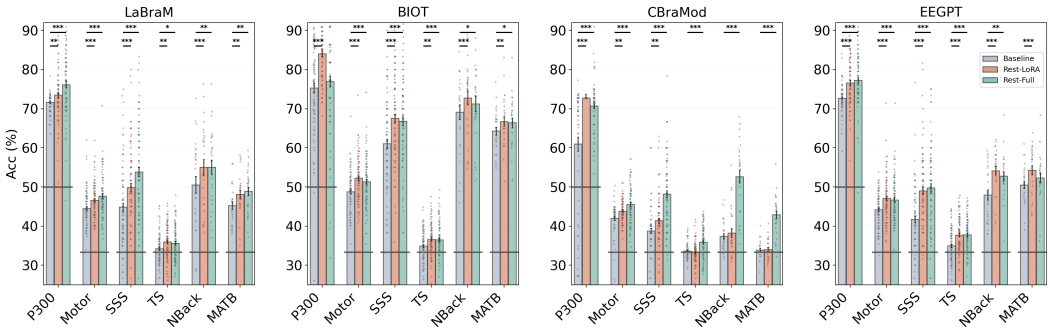

Figure 2: **Linear probing performance.** Bar plots show uni-task decoding accuracy across six downstream tasks and four EEG FMs. Gray: no-rest baseline; orange: Rest-LoRA; green: Rest-Full. Error bars denote standard deviation. Significance was tested with paired-subject comparisons (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

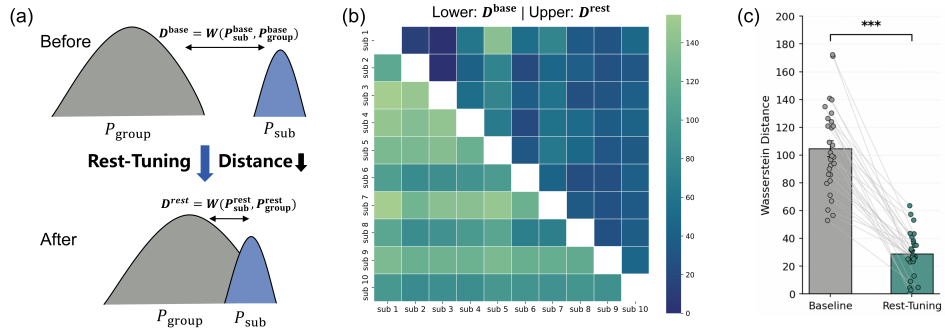

Figure 3: **Rest-Tuning reduces subject-to-group mismatch.** (a) Conceptual illustration: single-subject embedding (P_{sub}, blue) moves closer to the group prior (P_{group}, gray) after Rest-Tuning. (b) Pairwise Wasserstein distances for 10 subjects; lower triangle: baseline, upper triangle: Rest-Tuned. (c) Subject-to-group distances before (D_i^{base}) and after (D_i^{rest}) Rest-Tuning. Each dot represents a subject. Paired t-test: *** $p < 0.001$.

4.3 Rest-Tuning aligns subject representations

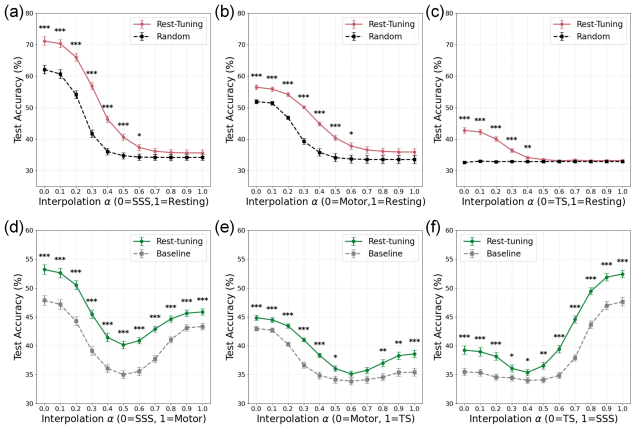

Figure 4: **LoRA adapter interpolation.** (a–c) Rest-to-task interpolation: Rest-Tuned adapters maintain higher accuracy than randomly initialized LoRA. (d–f) Task-to-task interpolation. Paired t-tests: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Rest-Tuning aligns subject-specific representations with the group prior. Pretrained EEG FMs capture a group-level representation prior, but individual subjects exhibit idiosyncratic neural and signal characteristics that shift their embeddings away from this prior. Rest-Tuning uses unlabeled resting EEG to reduce this subject-to-group mismatch before task-specific learning (Figure 3a).

We examine this alignment using LaBraM on COGBCI, constructing the group and subject embedding distributions following Sec. 3.4. Figure 3b compares pairwise Wasserstein distances among 10 subjects, with baseline distances in the lower triangle and Rest-Tuned distances in the upper triangle. Figure 3c quantifies the alignment with the group distribution: Rest-Tuning reduces the mean subject-to-group Wasserstein distance from 104.59 to 28.68

(paired t-test, $p < 0.001$). These results show that resting-state calibration brings subject-specific representations closer to the group distribution before any task supervision. We further confirm this reduction with CBraMod and BIOT on COGBCI and with LaBraM on M3CV (Appendix E).

Rest-Tuning improves adapter-space compatibility. Building on the reduction of subject-to-group mismatch, we examine how Rest-Tuning affects task-specific adaptation. By aligning embeddings across subjects, Rest-Tuning creates a smoother, more consistent parameter space, facilitating downstream adaptation. We probe this effect using LoRA adapters to explore the low-loss landscape. Rest-to-task interpolation (Figure 4a–c) shows Rest-Tuning adapters maintain higher accuracy than randomly initialized LoRA, placing them closer to task-relevant low-loss regions. Task-to-task interpolation (Figure 4d–f) reduces performance drops between adapters, demonstrating a smoother, more reliable adapter landscape that facilitates cross-task transfer. Additional trial-level t-SNE visualizations (Figure A5) show that Rest-Tuning enhances task-cluster separability across adaptation stages, indicating more stable and task-informative embeddings.

4.4 Rest-Tuning sustains personalization over time

We evaluate whether the benefits of a single resting calibration persist across recording days using COGBCI data from all 29 subjects, collected at an initial visit and one and two weeks later. For each subject, EEG FM is calibrated using only the initial resting recording and then frozen. At each time point, a separate linear head is trained using 10% of the available task labels. The task-only baseline applies the same linear-probing protocol to the original pretrained backbone with identical data splits.

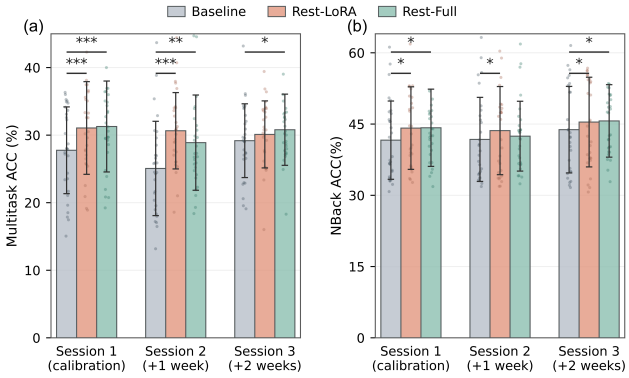

Figure 5: **Resting calibration benefits persist across weeks.** LaBraM performance on COGBCI for (a) multi-task and (b) NBack decoding at weeks 0, 1, and 2. Rest-LoRA and Rest-Full use initial-session rest only; backbones are then frozen, with session-specific linear heads trained on 10% of task labels. Nominal paired-test thresholds: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure 5 shows the LaBraM results for multi-task and NBack decoding. Rest-Full improves mean accuracy in both decoding settings at both follow-ups, including a 1.62 percentage-point gain in multi-task accuracy two weeks after calibration. These sustained gains show that resting calibration produces a personalized backbone whose benefits extend beyond the day of acquisition, supporting subsequent task learning without repeated resting calibration. Additional evaluations with CBraMod show the same pattern of Rest-Full improvements at both follow-ups (Table A10).

4.5 Rest-Tuning reduces supervision and calibration cost

One rest calibration reduces task-label dependence across tasks. We vary the labeled training budget to evaluate whether a retained resting calibration reduces the amount of supervised task data needed for a given accuracy.

Figure 6 shows that Rest-Tuned models reach baseline accuracy with substantially fewer labeled trials. In the multi-task setting of M3CV, Rest-LoRA and Rest-Full save up to 34.2 PP and 43.4 PP of labeled training data, respectively. Across individual tasks and datasets (M3CV and COGBCI), the reduction ranges from 20 PP to 60 PP, indicating that a single resting calibration can be reused across multiple task domains. Rest-Full generally provides larger savings, while Rest-LoRA offers a parameter-efficient alternative. These results suggest that Rest-Tuning does not replace task supervision, but substantially reduces task-label dependence by amortizing subject calibration through resting EEG. Since we only tested training fractions from 10% to 70%, the observed maximum savings is 60 pp. Additional model- and task-specific results are reported in Appendix J.

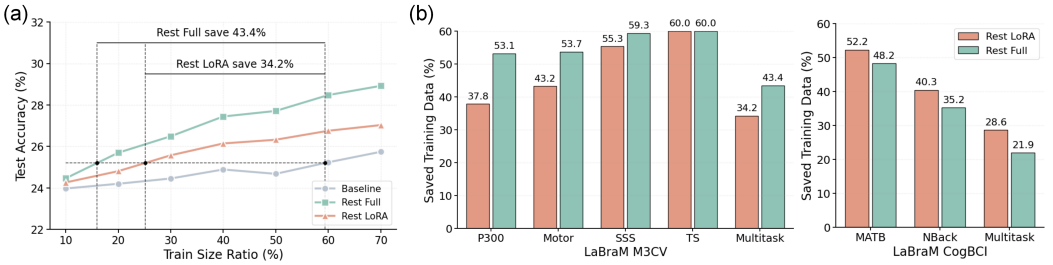

Figure 6: **Rest-Tuning reduces labeled task data requirements.** (a) Test accuracy of LaBraM on M3CV multitask under varying training sample ratios for baseline, Rest-LoRA, and Rest-Full. (b) Data-efficiency summaries across uni-task and multi-task evaluations in the two datasets.

Rest-Tuning requires only minutes of rest EEG. To assess the amount of resting-state EEG required for practical subject-specific adaptation, we varied the duration of EEG used during Rest-Tuning while keeping all other settings fixed. Figure 7 shows LaBraM Rest-Full; Table A14 extends the analysis to LaBraM, EEGPT, and BIOT with both calibration interfaces, averaging over task-label budgets from 1% to 10%. Accuracy improves rapidly with increasing rest data, reaching near-saturation after approximately 60–180 seconds, with only marginal gains beyond three minutes. These results indicate that 1–3 minutes of resting-state EEG are sufficient to effectively calibrate the subject-specific backbone for multi-task decoding, minimizing the data collection burden while maintaining strong downstream performance.

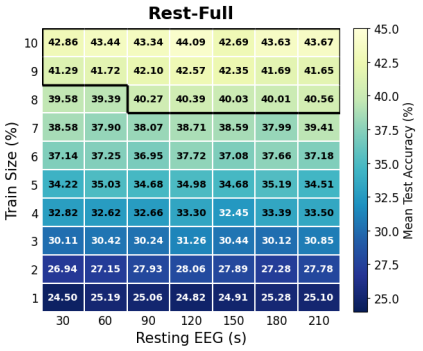

Figure 7: **Calibration duration and multi-task accuracy.** LaBraM Rest-Full across resting-recording durations and task-label budgets.

Computational Efficiency of Rest-Tuning We evaluate the computational overhead of Rest-Tuning on the M3CV dataset using 70% of training data. Across LaBraM and BIOT models, adapting a subject-specific backbone via Rest-Tuning requires under 3 minutes on average per subject (see Table A12 for detailed timings). Critically, Rest-Tuning is task-agnostic: a single calibration can be reused across multiple downstream tasks. As shown in Figure 8, the effective per-task cost decreases as the number of tasks increases. When adapting four tasks, the cost relative to Task-SFT drops to 38.3% for LaBraM and 51.8% for BIOT. These results indicate that Rest-Tuning introduces only a modest one-time computational overhead while providing increasing efficiency in multiple downstream tasks BCI scenarios, highlighting its practical applicability.

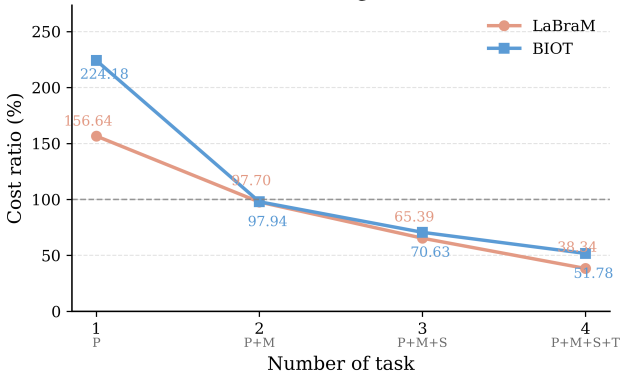

Figure 8: **Effective per-task cost of Rest-Tuning relative to Task-SFT on the M3CV dataset.** The four tasks are P: P300, M: Motor, S: Steady-State Sensory, and T: Transient Sensory.

Accuracy–cost trade-off across lightweight and foundation backbones. Finally, we compare task-only EEGNet and EEGConformer with Rest-Tuned LaBraM and BIOT under the same multi-task splits. The FMs use Rest-Full calibration followed by Task-Full adaptation. Table 2 reports accuracy at four train-size ratios and adaptation cost at the 70% label setting. Rest-Tuned FMs achieve substantially higher accuracy, particularly in low-label regimes, at comparable or higher adaptation cost. Resting calibration adds a one-time cost per subject, and the resulting backbone can be reused across downstream tasks, amortizing this overhead over repeated use. Full per-ratio cost and accuracy results are provided in Table A11.

Table 2: **Backbone accuracy–cost context for multitask decoding.** ACC is reported as mean accuracy (%) across subjects at different labeled train ratios. Cost@70% is the task-adaptation time in seconds at the 70% label setting and includes the Rest-Tuning stage. Lightweight decoders are trained directly on task labels, whereas FM rows use Rest-Full followed by Task-Full adaptation.

Dataset	Model	Protocol	Rest-Tuning Cost (s)	ACC@5%	ACC@10%	ACC@50%	ACC@70%	Cost@70% (s)
COGBCI	EEGNet	Task-only	/	21.47	22.10	62.70	64.38	**287.80**
	EEGConformer	Task-only	/	28.93	34.00	52.77	59.54	1045.37
	LaBraM	Rest+Task-Full	150.37	35.67	44.42	72.85	78.13	528.71
	BIOT	Rest+Task-Full	80.90	**70.41**	**78.56**	**89.32**	**90.45**	369.66
M3CV	EEGNet	Task-only	/	22.08	24.69	34.92	39.38	757.33
	EEGConformer	Task-only	/	24.42	24.14	28.59	32.59	764.47
	LaBraM	Rest+Task-Full	128.67	25.36	28.08	44.08	49.79	**251.23**
	BIOT	Rest+Task-Full	69.72	**30.28**	**36.35**	**47.91**	**52.82**	198.10

5 Discussion and Limitations

We introduced Rest-Tuning, a resting-state calibration interface for adapting EEG FMs to individuals. Our central message is that EEG FMs provide useful population priors, but require a low-cost subject interface for reliable subject-level adaptation. Rest-Tuning provides this interface by using resting EEG to obtain a reusable subject-calibrated backbone before task-specific supervision.

Reusable subject calibration. Rest-Tuning separates subject calibration from task learning. Unlabeled target-subject resting EEG produces a personalized backbone shared across downstream tasks, while task adaptation uses the same labeled-data budgets as task-only baselines. This differs from BCM and ResTL, which learn task-specific models using source-subject resting and labeled task recordings. Rest-Tuning draws on an existing population FM to calibrate the subject representation without learning an explicit rest–task mapping (Appendix H). The resulting checkpoint retains the subject-specific adjustment and makes it available to subsequent tasks.

Representation alignment and transfer. The representation analyses connect the decoding gains to changes in both embedding and adapter geometry. Reduced subject-to-group Wasserstein distances indicate closer alignment between individual and population representations. Smoother LoRA interpolation paths further suggest that resting calibration produces an initialization compatible with different supervised task solutions. The t-SNE visualizations provide a complementary view of clearer within-subject task separation (Appendix L). The loss ablations and matched MAE comparison show that the calibration objective matters for downstream transfer, supporting the combination of contrastive discrimination and matched-pair consistency against pretrained representations (Appendix C). Together, these results support teacher-guided representation alignment as an effective way to extract transferable subject information from resting EEG.

Persistence and practical value. The benefits of retaining a subject calibration extend beyond reuse across tasks. On COGBCI, the initial resting calibration continues to improve linear probing one and two weeks later. This demonstrates that the calibrated representation remains useful beyond the recording day on which it was learned. Combined with calibration from 1–3 minutes of resting EEG and reduced task-label requirements, this persistence makes the initial acquisition and computation useful across subsequent tasks and visits. The lightweight-backbone comparison places this benefit within an accuracy–cost trade-off: Rest-Tuned FMs achieve higher decoding accuracy, particularly with limited labels, at comparable or higher adaptation cost. Their practical advantage is therefore strongest when accurate decoding and repeated use of the personalized backbone are both important.

Limitations and future directions. Rest-Tuning currently retains a static calibration, and its temporal evaluation covers two weeks under the COGBCI acquisition protocol. An important next step is to determine when calibration should be refreshed as neural state, electrode contact, or recording conditions change. Resting artifacts and differences between rest and task acquisition also motivate quality-aware calibration and periodic or online updates.

Acknowledgments and Disclosure of Funding

This work was supported by National Natural Science Foundation of China (3254100307, 62472206), Brain Science and Brain-like Intelligence Technology-National Science and Technology Major Project (2021ZD0200500, 2026ZD0220402), National Key R&D Program of China (2025YFC3410000), Shenzhen Science and Technology Innovation Committee (RCYX20231211090405003), Guangdong Basic and Applied Basic Research Foundation (2026B1515020099), Guangdong S&T Program (2026B0101110003), Shanghai Municipal Special Program for Basic Research on General AI Foundation Models (2025SHZDZX026D05), Shenzhen Loop Area Institute (FPF10120250012), and the open research fund of the Guangdong Provincial Key Laboratory of Mathematical and Neural Dynamical Systems, the Center for Computational Science and Engineering at Southern University of Science and Technology, Shenzhen Key Laboratory of Smart Healthcare Engineering.

References

[1] Bruno Aristimunha, Dung Truong, Pierre Guetschel, Seyed Yahya Shirazi, Isabelle Guyon, Alexandre R Franco, Michael P Milham, Aviv Dotan, Scott Makeig, Alexandre Gramfort, et al. Eeg foundation challenge: from cross-task to cross-subject eeg decoding. *arXiv preprint arXiv:2506.19141*, 2025.

[2] Dingkun Liu, Yuheng Chen, Zhu Chen, Zhenyao Cui, Yaozhi Wen, Jiayu An, Jingwei Luo, and Dongrui Wu. Eeg foundation models: Progresses, benchmarking, and open problems. *arXiv preprint arXiv:2601.17883*, 2026.

[3] Erin Gibson, Nancy J Lobaugh, Steve Joordens, and Anthony R McIntosh. Eeg variability: Task-driven or subject-driven signal of interest? *NeuroImage*, 252:119034, 2022.

[4] Liuyin Yang, Qiang Sun, Ang Li, and Marc M. Van Hulle. Are EEG foundation models worth it? comparative evaluation with traditional decoders in diverse BCI tasks. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=5Xwm8e6vbh.

[5] Timon Klein, Piotr Minakowski, and Sebastian Sager. Mitigating subject dependency in EEG decoding with subject-specific low-rank adapters. In *NeurIPS 2025 Workshop on Foundation Models for the Brain and Body*, 2025. URL https://openreview.net/forum?id=vDtsuVfBmG.

[6] Chenrui Wu, Haishuai Wang, Xiang Zhang, Chengqi Zhang, and Jiajun Bu. Efficient personalized adaptation for physiological signal foundation model. In *Forty-second International Conference on Machine Learning*, 2025. URL https://openreview.net/forum?id=55ysNwbOTI.

[7] Özgün Turgut, Felix S Bott, Markus Ploner, and Daniel Rueckert. Are foundation models useful feature extractors for electroencephalography analysis? *arXiv preprint arXiv:2502.21086*, 2025.

[8] Wei-Bang Jiang, Yansen Wang, Bao-Liang Lu, and Dongsheng Li. NeuroLM: A universal multi-task foundation model for bridging the gap between language and EEG signals. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=Io9yFt7XH7.

[9] Gayal Kuruppu, Neeraj Wagh, and Yogatheesan Varatharajah. EEG foundation models: A critical review of current progress and future directions. In *NeurIPS 2025 Workshop on Foundation Models for the Brain and Body*, 2025. URL https://openreview.net/forum?id=Iu6qVgtgUD.

[10] Weibang Jiang, Liming Zhao, and Bao liang Lu. Large brain model for learning generic representations with tremendous EEG data in BCI. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=QzTpTRVtrP.

[11] Chaoqi Yang, M Brandon Westover, and Jimeng Sun. BIOT: Biosignal transformer for cross-data learning in the wild. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023. URL https://openreview.net/forum?id=c2LZyTyddi.

[12] Guangyu Wang, Wenchao Liu, Yuhong He, Cong Xu, Lin Ma, and Haifeng Li. EEGPT: Pretrained transformer for universal and reliable representation of EEG signals. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*, 2024. URL https://openreview.net/forum?id=lvS2b8CjG5.

[13] Jiquan Wang, Sha Zhao, Zhiling Luo, Yangxuan Zhou, Haiteng Jiang, Shijian Li, Tao Li, and Gang Pan. CBramod: A criss-cross brain foundation model for EEG decoding. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=NPNUHgHF2w.

[14] Takuya Ito, Kaustubh R Kulkarni, Douglas H Schultz, Ravi D Mill, Richard H Chen, Levi I Solomyak, and Michael W Cole. Cognitive task information is transferred between brain regions via resting-state network topology. *Nature communications*, 8(1):1027, 2017.

[15] Michael W Cole, Takuya Ito, Danielle S Bassett, and Douglas H Schultz. Activity flow over resting-state networks shapes cognitive task activations. *Nature neuroscience*, 19(12): 1718–1726, 2016.

[16] Luyao Wang, Jian Zhang, Tiantian Liu, Duanduan Chen, Dikun Yang, Ritsu Go, Jinglong Wu, and Tianyi Yan. Prediction of cognitive task activations via resting-state functional connectivity networks: an eeg study. *IEEE Transactions on Cognitive and Developmental Systems*, 14(1): 181–188, 2020.

[17] Minji Lee, Jae-Geun Yoon, and Seong-Whan Lee. Predicting motor imagery performance from resting-state eeg using dynamic causal modeling. *Frontiers in human neuroscience*, 14:321, 2020.

[18] Lan Ma, James W Minett, Thierry Blu, and William SY Wang. Resting state eeg-based biometrics for individual identification using convolutional neural networks. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2848–2851. IEEE, 2015.

[19] Youngchul Kwak, Kyeongbo Kong, Woo-Jin Song, and Seong-Eun Kim. Subject-invariant deep neural networks based on baseline correction for eeg motor imagery bci. *IEEE Journal of Biomedical and Health Informatics*, 27(4):1801–1812, 2023.

[20] Sion An, Myeongkyun Kang, Soopil Kim, Philip Chikontwe, Li Shen, and Sang Hyun Park. Subject-adaptive transfer learning using resting state eeg signals for cross-subject eeg motor imagery classification. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 678–688. Springer, 2024.

[21] Sion An, Myeongkyun Kang, Soopil Kim, Philip Chikontwe, Li Shen, and Sang Hyun Park. Subject-adaptive meta-learning for personalized bci: A fusion of resting-state eeg signal and task-specific information. *Information Fusion*, page 103501, 2025.

[22] Jingying Ma, Feng Wu, Qika Lin, Yucheng Xing, Chenyu Liu, Ziyu Jia, and Mengling Feng. CodeBrain: Bridging decoupled tokenizer and multi-scale architecture for EEG foundation model. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=msJgEkjwh5.

[23] Demetres Kostas, Stéphane Aroca-Ouellette, and Frank Rudzicz. BENDR: Using transformers and a contrastive self-supervised learning task to learn from massive amounts of EEG data. *Frontiers in Human Neuroscience*, 15:653659, 2021. doi: 10.3389/fnhum.2021.653659.

[24] Hsiang-Yun Sherry Chien, Hanlin Goh, Christopher M. Sandino, and Joseph Y. Cheng. MAEEG: Masked auto-encoder for EEG representation learning. In *NeurIPS Workshop on Learning from Time Series for Health*, 2022. URL https://arxiv.org/abs/2211.02625.

[25] Navid Mohammadi Foumani, Geoffrey Mackellar, Soheila Ghane, Saad Irtza, Nam Nguyen, and Mahsa Salehi. EEG2Rep: Enhancing self-supervised EEG representation through informative masked inputs. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 5544–5555. Association for Computing Machinery, 2024. doi: 10.1145/3637528.3671600.

[26] Kun Yi, Yixiao Ge, Xiaotong Li, Shusheng Yang, Dian Li, Jianping Wu, Ying Shan, and Xiaohu Qie. Masked image modeling with denoising contrast. In *International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=1fZd4owfJP6.

[27] Zhicheng Huang, Xiaojie Jin, Chengze Lu, Qibin Hou, Ming-Ming Cheng, Dongmei Fu, Xiaohui Shen, and Jiashi Feng. Contrastive masked autoencoders are stronger vision learners. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(4):2506–2517, 2024. doi: 10.1109/TPAMI.2023.3336525.

[28] Han Wei Ng and Cuntai Guan. Subject-independent meta-learning framework towards optimal training of eeg-based classifiers. *Neural Networks*, 172:106108, 2024.

[29] Jingying Ma, Feng Wu, Yucheng Xing, Qika Lin, Tianyu Liu, Chenyu Liu, Ziyu Jia, and Mengling Feng. SCOPE: Structured prototype-guided adaptation for EEG foundation models with limited labels. *arXiv preprint arXiv:2602.17251*, 2026. URL https://arxiv.org/abs/2602.17251.

[30] Suli Wang, Yangshen Deng, Zhenghua Bao, Xinyu Zhan, and Yiqun Duan. Neurottt: Bridging pretraining-downstream task misalignment in eeg foundation models via test-time training, 2025. URL https://arxiv.org/abs/2509.26301.

[31] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*, pages 1597–1607. PmLR, 2020.

[32] Kaiming He, Xinlei Chen, Saining Xie, Yanghao Li, Piotr Dollár, and Ross Girshick. Masked autoencoders are scalable vision learners. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 16000–16009, 2022.

[33] Aaron van den Oord, Yazhe Li, and Oriol Vinyals. Representation learning with contrastive predictive coding. *arXiv preprint arXiv:1807.03748*, 2018.

[34] Jonathan Frankle, Gintare Karolina Dziugaite, Daniel Roy, and Michael Carbin. Linear mode connectivity and the lottery ticket hypothesis. In *International Conference on Machine Learning*, pages 3259–3269. PMLR, 2020.

[35] Junsu Kim, Jaeyeon Kim, and Ernest K Ryu. Lora training provably converges to a low-rank global minimum or it fails loudly (but it probably won't fail). *arXiv preprint arXiv:2502.09376*, 2025.

[36] Tao Li, Zhengbao He, Yujun Li, Yasheng Wang, Lifeng Shang, and Xiaolin Huang. Flat-lora: Low-rank adaptation over a flat loss landscape. *arXiv preprint arXiv:2409.14396*, 2024.

[37] Gan Huang, Zhenxing Hu, Weize Chen, Shaorong Zhang, Zhen Liang, Linling Li, Li Zhang, and Zhiguo Zhang. M3cv: A multi-subject, multi-session, and multi-task database for eeg-based biometrics challenge. *NeuroImage*, 264:119666, 2022.

[38] Marcel F Hinss, Emilie S Jahanpour, Bertille Somon, Lou Pluchon, Frédéric Dehais, and Raphaëlle N Roy. Open multi-session and multi-task eeg cognitive dataset for passive brain-computer interface applications. *Scientific Data*, 10(1):85, 2023.

A Datasets and Data Processing

M3CV M3CV is a multi-subject, multi-session, and multi-task EEG benchmark comprising 106 subjects recorded in two sessions, with 15 runs per subject, covering six paradigms and 14 tasks (approximately $\sim$120,000 epochs in total) [37]. Recordings use a 64-channel EEG setup following the 10–20 system. The paradigms span (i) *resting state* with eyes-open and eyes-closed conditions, (ii) *transient sensory* evoked responses (visual, auditory, and somatosensory), (iii) *steady-state sensory* responses (SSVEP, SSAEP, and SSSEP), (iv) a *visual oddball* paradigm eliciting P300 responses (target vs. non-target), and (v) *motor execution* tasks (e.g., foot, right-hand, and left-hand movements). This wide coverage enables unified evaluation of models across heterogeneous sensory, cognitive, and motor conditions, as well as across recording sessions.

COGBCI COGBCI is an open multi-session, multi-task EEG dataset targeting passive BCI applications, with 29 participants recorded over three sessions spaced one week apart, totaling over 100 hours of EEG [38]. Each session includes (i) *resting state* (eyes-open and eyes-closed) recorded at both the beginning and end of the protocol, and (ii) four cognitive tasks designed to elicit distinct mental states: a *Psychomotor Vigilance Task (PVT)* for vigilance/fatigue, an *N-back* task (0/1/2-back) for graded workload, an arrow-based *Flanker* task for conflict/error processing, and *MATB-II* for more ecological multi-attribute workload manipulation (easy/medium/difficult). EEG is acquired with a 64-channel active-electrode system at 500 Hz (10–20 layout), and the released data are provided in BIDS format with behavioral and subjective measures (e.g., KSS/RSME) for validation and multimodal modeling.

Data Preprocessing To standardize the heterogeneous recordings from M3CV and COGBCI with the input specifications of the EEG FMs, we applied a consistent preprocessing pipeline. Let the raw continuous EEG signal be denoted as $\mathbf{D}_{\text{raw}} \in \mathbb{R}^{C \times L_{\text{raw}}}$, where C represents the channel dimension and f_s the original sampling rate. We first downsampled all signals to a target rate of $f_s' = 200$ Hz and segmented them into fixed-length, non-overlapping windows of duration $T = 1$ second. This yields model inputs $\mathbf{X} \in \mathbb{R}^{C \times N}$, where $N = 200$. To mitigate inter-subject amplitude variability, we applied instance-wise Z-score standardization. For label space construction, we formulated two strategies: for single-task analysis, we preserved the intrinsic local labels y_{local}; for the multi-task setting, to resolve label collisions across different paradigms, we constructed a global label space $\mathcal{Y}_{\text{global}}$ using a strided mapping function Φ. Given a task index i and a stride $\delta > \max_i(K_i)$, the global label is defined as:

$$y_{\text{global}} = \Phi(i, y_{\text{local}}) = i \times \delta + y_{\text{local}}, \tag{12}$$

which projects all task conditions into a unified categorical space, enabling simultaneous evaluation of the personalized backbone across disjoint tasks.

Table A1: Summary of datasets, sessions, tasks, classes, and average numbers of segments per subject.

Dataset	Subjects	Session	Resting Segments (mean)	Task	Classes	Task Segments (mean)
COGBCI	29	1	240.0	MATB	3	897.6
				NBack	3	1042.1
		2	/	MATB	3	897.0
				NBack	3	981.4
		3	/	MATB	3	897.0
				NBack	3	964.4
M3CV	95	1	265.2	Motor	3	814.6
				P300	2	486.5
				Steady-State Sensory	3	332.4
				Transient Sensory	3	198.1

B Foundation Model Summary and Parameter Count

Table A2 summarizes the characteristics of the EEG foundation models used in this work, including model scale, architectural design, pretraining strategy, pretraining signal duration, and supported EEG channels. The four models differ substantially in pretraining scale and channel configurations, ranging from 198 hours to 59,304 hours of pretraining EEG signals and supporting different electrode settings.

These differences provide a diverse benchmark for evaluating the effectiveness and generalizability of the proposed Rest-Tuning framework across heterogeneous EEG foundation models.

Table A2: **Summary of recently published brain foundation models.**

Model	Parameters	Architecture / Pretraining Strategy	Pretraining Signal Duration	Supported Channels
BIOT [11]	3.2M	Linear-Transformer / Self-supervised - Contrastive Learning	59,304h	18
EEGPT [12]	25M	CNN-Transformer / Masked Signal Reconstruction	198h	58
LaBraM [10]	5.8M	CNN-Transformer / Masked Signal Reconstruction	2,535h	136
CBraMod [13]	4.0M	CNN-Transformer / Masked Signal Reconstruction	9,246h	Montage-agnostic

Table A3 reports the total and trainable parameters under Full tuning, LoRA tuning, and linear probing. For Full and LoRA tuning, the trainable ratio is calculated with respect to the total parameters of the corresponding tuning strategy. Compared with Full tuning, LoRA reduces the trainable parameters to only 3.03%–5.00%, while linear probing requires only 1,206 trainable parameters for each model.

Table A3: **Total and trainable parameters under different tuning strategies.**

Model	Strategy	Reference Total Params	Trainable Params	Ratio
LaBraM	Full tuning	6,017,936	6,017,936	100.00%
	LoRA tuning	6,334,736	316,800	5.00%
	Linear probing	6,017,936	1,206	0.020%
BIOT	Full tuning	3,186,706	3,186,688	100.00%
	LoRA tuning	3,337,018	150,312	4.50%
	Linear probing	3,186,706	1,206	0.038%
CBraMod	Full tuning	4,924,000	4,924,000	100.00%
	LoRA tuning	5,077,600	153,600	3.03%
	Linear probing	4,924,000	1,206	0.024%
EEGPT	Full tuning	25,287,168	25,287,168	100.00%
	LoRA tuning	25,811,456	524,288	2.03%
	Linear probing	25,287,168	1,206	0.005%

C Model Hyperparameters and Ablation Studies

Table A4 summarizes the hyperparameter settings used for Rest-Tuning. We examine the contributions of the calibration loss terms, sensitivity to the mask ratio, and robustness to the MSE weight below. Unless otherwise stated, results are multitask classification accuracies (%), reported as mean $\pm$ standard deviation across target subjects. Rest-Tuning and task-only baselines use identical downstream data splits and adaptation interfaces.

C.1 Loss-Term Ablation

We evaluate the contributions of the contrastive and MSE terms using LaBraM on all 29 COGBCI subjects, with Rest-Full calibration followed by Task-Full adaptation. Table A5 compares the complete objective with either term removed, as well as a task-only baseline without resting calibration, under 10% and 70% labeled training budgets. Two-sided paired t-tests compare subject-level accuracies between each alternative and the complete objective.

C.2 Comparison with MAE Calibration

We compare Rest-Tuning with masked autoencoder (MAE) calibration using LaBraM on all 29 COGBCI subjects under the multitask setting with 10% labeled training data. Both methods use the same target-subject resting EEG, calibration parameterization, and downstream adaptation interface.

Table A4: Key hyperparameters for Rest-Tuning and their rationale.

Parameter	Value	Update Strategy	Rationale
Mask ratio ρ	0.3(Full)/0.7(LoRA)	Fixed	Calibration-specific defaults evaluated in the mask-ratio sensitivity analysis.
MSE weight λ_{mse}	0.1	Fixed	Ensures embedding space smoothness and amplifies positive-pair gradient.
Temperature τ	$0.1 \rightarrow 0.005$	Cosine annealing	High initial temperature stabilizes early representations; decayed temperature sharpens inter-instance discrimination.
Batch size n	32	Fixed	Ensures sufficient negative samples for contrastive learning; suitable for small resting EEG datasets.
Epoch number	Max 300, early stop patience=30	Early stopping	Prevents overfitting on short resting EEG recordings.
Learning rate	1e-6 $\rightarrow$ 1e-4	Linear warm-up + cosine decay	Warm-up stabilizes early training; cosine decay smooths convergence to final solution.
Optimizer	AdamW	Fixed	Decouples weight decay from gradient update.
LoRA rank	8	Fixed	Low-rank adaptation balances parameter efficiency and expressive capacity.
LoRA alpha	32	Fixed	Scales the residual update to control magnitude of LoRA adaptation.
LoRA dropout	0.1	Fixed	Regularizes LoRA adapters to reduce overfitting.

Table A5: Loss-term ablation on COGBCI using LaBraM with Rest-Full + Task-Full. Accuracies are mean $\pm$ standard deviation across 29 subjects. p-values compare each row with the complete objective using paired subject-level tests.

Calibration objective	InfoNCE	MSE	Train 10%		Train 70%	
			ACC (%)	p vs. complete	ACC (%)	p vs. complete
No Rest-Tuning			39.41 ± 7.44	5.6×10^{-7}	73.44 ± 9.53	3.2×10^{-18}
MSE only		✓	42.28 ± 9.27	3.8×10^{-3}	75.80 ± 9.71	1.9×10^{-2}
InfoNCE only	✓		41.31 ± 8.72	3.5×10^{-3}	77.03 ± 10.07	0.15
Complete objective	✓	✓	$\mathbf{44.42 \pm 9.82}$	–	$\mathbf{78.13 \pm 9.45}$	–

MAE-LoRA and Rest-LoRA are followed by Task-LoRA, while MAE-Full and Rest-Full are followed by Task-Full. Table A6 reports accuracy and paired subject-level comparisons.

Table A6: Comparison with MAE calibration on COGBCI using LaBraM with 10% labeled training data. Accuracy is reported as mean $\pm$ standard deviation across 29 subjects. p-values compare MAE with Rest-Tuning under the corresponding calibration and task adaptation interfaces.

Calibration	Task adaptation	ACC (%)	Paired p
MAE-LoRA	Task-LoRA	37.42 ± 6.74	2.76×10^{-5}
Rest-LoRA	Task-LoRA	$\mathbf{39.78 \pm 7.24}$	–
MAE-Full	Task-Full	38.46 ± 9.20	1.95×10^{-4}
Rest-Full	Task-Full	$\mathbf{44.42 \pm 9.82}$	–

C.3 Mask-Ratio Sensitivity

We vary the mask ratio $\rho \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ using LaBraM on COGBCI at the 10% labeled training budget. We evaluate Rest-Full + Task-Full and Rest-LoRA + Task-LoRA separately, keeping the remaining settings fixed.

Table A7: Mask-ratio sensitivity on COGBCI using LaBraM with 10% labeled training data. Bold indicates the highest mean within each calibration interface.

Mask ratio ρ	Rest-Full + Task-Full	Rest-LoRA + Task-LoRA
No Rest-Tuning	39.41 ± 7.44	37.50 ± 6.13
0.1	43.16 ± 9.72	38.45 ± 6.79
0.3	$\mathbf{44.42 \pm 9.82}$	39.41 ± 6.84
0.5	43.41 ± 9.42	39.02 ± 6.95
0.7	43.17 ± 9.52	$\mathbf{39.78 \pm 7.24}$
0.9	42.22 ± 9.15	39.01 ± 7.32

As shown in Table A7, all tested mask ratios yield higher mean accuracy than their corresponding task-only baselines. Rest-Full and Rest-LoRA attain their highest means at $\rho = 0.3$ and $\rho = 0.7$, respectively, supporting the defaults used in our experiments. The mean accuracy spans 2.20 pp for Rest-Full and 1.33 pp for Rest-LoRA across the tested ratios. These results indicate interface-dependent sensitivity; they do not establish that the selected ratios are optimal for every backbone or dataset.

C.4 MSE-Weight Sensitivity

To examine whether the default MSE weight transfers beyond a single backbone–dataset combination, we evaluate $\lambda_{\text{mse}} \in \{0.01, 0.1, 1\}$ for LaBraM and CBraMod on COGBCI, and LaBraM on M3CV, under both 10% and 70% labeled training budgets. All settings use Rest-Full followed by Task-Full. The $\lambda_{\text{mse}} = 0$ condition for LaBraM–COGBCI is reported as InfoNCE only in Table A5.

Table A8: Sensitivity to the MSE weight under Rest-Full + Task-Full. Results are multitask accuracies (%). Bold indicates the highest mean in each row.

Dataset	Backbone	Training budget	λ_{mse}		
			0.01	0.1	1
COGBCI	LaBraM	10%	43.91 ± 9.83	$\mathbf{44.42 \pm 9.82}$	42.99 ± 9.15
		70%	77.88 ± 9.01	$\mathbf{78.13 \pm 9.45}$	77.28 ± 9.46
	CBraMod	10%	49.63 ± 9.94	$\mathbf{50.17 \pm 8.84}$	49.43 ± 9.26
		70%	71.71 ± 9.47	73.69 ± 8.62	$\mathbf{73.72 \pm 8.56}$
M3CV	LaBraM	10%	28.57 ± 4.13	$\mathbf{28.08 \pm 3.85}$	27.96 ± 4.03
		70%	49.29 ± 7.55	$\mathbf{49.79 \pm 7.71}$	49.60 ± 7.43

Across this 100-fold weight range, the within-setting variation in mean accuracy is at most 2.01 pp. The default $\lambda_{\text{mse}} = 0.1$ achieves the highest mean in four of the six settings and is only 0.03 pp below the highest mean in the remaining setting. These results support $\lambda_{\text{mse}} = 0.1$ as a stable default across the evaluated combinations, rather than establishing its statistical superiority over alternative weights.

D Teacher-Student Architecture for Rest-Tuning

Figure A1 illustrates the teacher-student architecture used in Rest-Tuning. Resting-state EEG segments are divided into temporal windows $t_1, t_2, \ldots, t_n$ and optionally masked to create $\tilde{x}_i^r$ for the student network. The **student network** is trainable and receives masked EEG inputs, learning to

encode features that are robust to missing channels or temporal patches. Simultaneously, the **teacher network** is frozen and encodes the unmasked EEG, providing a stable reference representing the population prior.

Student embeddings $z_i^{(st)}$ are aligned with teacher embeddings $z_i^{(te)}$ via a contrastive InfoNCE loss, ensuring that each masked student embedding is pulled towards its corresponding teacher embedding while being repelled from other embeddings in the mini-batch. Additionally, a mean-squared error consistency regularizer $\mathcal{L}_{\text{mse}}$ is applied to further smooth the embedding space and stabilize training.

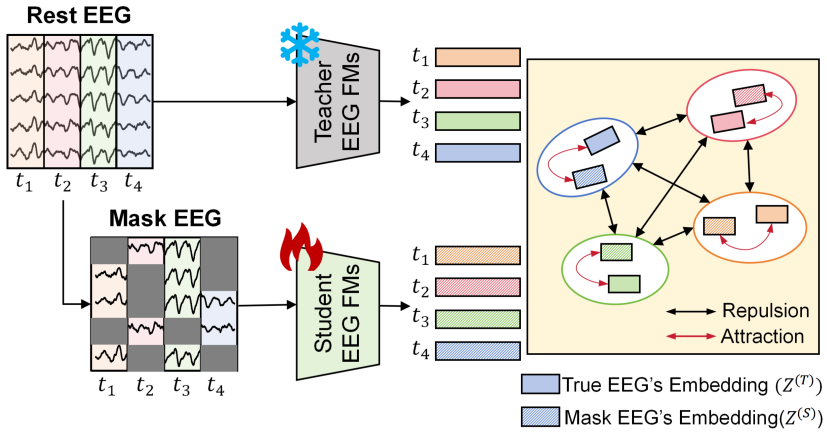

Figure A1: Illustration of the Rest-Tuning teacher-student architecture. Rest EEG segments are masked and fed into the trainable student network, while the frozen teacher network encodes the unmasked EEG. The student embeddings are aligned with the teacher embeddings via a contrastive loss and a consistency regularizer. This figure clarifies how masked EEG views and true EEG embeddings interact during calibration.

E Representation Alignment across Backbones and Datasets

To examine whether the representation-alignment effect extends beyond LaBraM on COGBCI, we additionally evaluate CBraMod and BIOT on COGBCI and LaBraM on M3CV. We follow the leave-one-subject-out procedure described in Section 3.4, including all 29 COGBCI subjects and 95 M3CV subjects. For each held-out subject, we construct a subject-excluded population model using the remaining subjects and then calibrate it using only the held-out subject's unlabeled resting EEG. Before and after calibration, target-subject and population embeddings are extracted using the same encoder within each condition, and their mismatch is quantified by the Wasserstein distance.

We compare pre/post distances using paired subject-level tests and report paired Cohen's d_z. For subject i, let $\Delta D_i = D_i^{\text{rest}} - D_i^{\text{base}}$; the effect size is $d_z = \text{mean}(\Delta D_i)/\,\text{SD}(\Delta D_i)$. Thus, negative values indicate a reduction in subject-to-group distance.

Table A9 shows significant paired reductions in Wasserstein distance for all four evaluated backbone–dataset combinations, with d_z ranging from -3.89 to -1.88. The additional results extend the alignment evidence to three backbones and both datasets, supporting that the observed reduction is not specific to LaBraM on COGBCI.

F Cross-Session Reuse of Rest-Tuned Backbones

We evaluate cross-session reuse on all 29 COGBCI subjects using LaBraM and CBraMod. For each subject, we perform Rest-Tuning once using Session-1 resting EEG and retain the calibrated backbone for evaluation in all three sessions. Sessions 2 and 3 were recorded one and two weeks after Session 1, respectively. The backbone remains frozen throughout downstream evaluation, with no additional resting calibration in later sessions.

Table A9: Subject-to-group Wasserstein distances before and after Rest-Tuning under the leave-one-subject-out analysis. Values are reported as mean ± standard error of the mean (SEM). Paired t-tests compare subject-level pre/post distances. Cohen's d_z is calculated using the sample standard deviation of paired differences; negative values indicate reduced distance.

Backbone	Dataset	n	Before Rest-Tuning	After Rest-Tuning	Paired p	Cohen's d_z
LaBraM	COGBCI	29	104.59 ± 5.74	28.68 ± 2.91	9.07×10^{-13}	-2.28
CBraMod	COGBCI	29	798.16 ± 28.72	176.81 ± 4.90	1.19×10^{-18}	-3.89
BIOT	COGBCI	29	163.04 ± 6.53	92.47 ± 2.86	7.41×10^{-11}	-1.88
LaBraM	M3CV	95	113.37 ± 4.39	8.53 ± 1.93	1.83×10^{-57}	-3.76

Within each session, we independently train a linear head using 10% of that session's task labels and evaluate it on the corresponding held-out test set. The task-only baseline uses the original pretrained backbone with identical task-data splits and linear-probing settings. We evaluate both merged-label multitask decoding and single-task NBack decoding. Table A10 reports subject-level accuracy statistics and paired comparisons with the session-matched task-only baseline.

Table A10: Cross-session linear-probing accuracy (%) on COGBCI. Rest-Tuning uses Session-1 resting EEG only; a separate linear head is trained with 10% of each session's task labels. Values are mean ± standard deviation across 29 subjects. † denotes nominal $p < 0.05$ in a paired subject-level t-test against the task-only baseline within the same backbone, decoding setting, and session.

Backbone	Setting	Initialization	S1	S2 (+1 week)	S3 (+2 weeks)
LaBraM	Multitask	Task only	27.76 ± 6.42	27.07 ± 6.98	29.18 ± 5.46
		Rest-LoRA	$31.06 \pm 6.86^{\dagger}$	$30.65 \pm 5.65^{\dagger}$	30.11 ± 4.96
		Rest-Full	$31.28 \pm 6.73^{\dagger}$	$28.89 \pm 7.05^{\dagger}$	$30.80 \pm 5.27^{\dagger}$
	NBack	Task only	41.61 ± 8.24	41.77 ± 8.85	43.82 ± 9.11
		Rest-LoRA	$44.15 \pm 8.69^{\dagger}$	$43.62 \pm 9.28^{\dagger}$	$45.42 \pm 9.44^{\dagger}$
		Rest-Full	$44.22 \pm 8.13^{\dagger}$	42.45 ± 7.35	$45.66 \pm 7.63^{\dagger}$
CBraMod	Multitask	Task only	19.53 ± 2.48	18.97 ± 3.84	19.04 ± 2.68
		Rest-LoRA	$22.33 \pm 4.05^{\dagger}$	18.46 ± 2.00	18.95 ± 2.28
		Rest-Full	$26.00 \pm 6.03^{\dagger}$	$24.24 \pm 6.20^{\dagger}$	$25.08 \pm 5.08^{\dagger}$
	NBack	Task only	33.89 ± 2.71	34.83 ± 3.74	33.56 ± 2.04
		Rest-LoRA	35.86 ± 5.51	35.15 ± 3.11	34.66 ± 3.39
		Rest-Full	$46.75 \pm 9.47^{\dagger}$	$39.89 \pm 7.38^{\dagger}$	$41.38 \pm 6.85^{\dagger}$

G Lightweight-backbone accuracy-cost comparison

We include supplementary tables for the lightweight-backbone accuracy-cost comparison A11.

H Comparison with Other Resting-State Calibration Methods

We compare Rest-Tuning with BCM [19] and ResTL [20] on all 29 COGBCI subjects under the multitask setting. BCM and ResTL use rest–task pairs from the other 28 subjects and target-subject resting EEG, without labeled target-task trials. Rest-Tuning uses only target-subject resting EEG for Rest-Full calibration, followed by Task-Full adaptation with 10% or 70% of the target-task labels. Table A13 reports training time and test accuracy averaged over the 29 target subjects.

Table A11: Full accuracy and cost comparison for multitask decoding. Cost is reported in seconds as mean ± standard deviation. ACC is reported as mean ± standard deviation across subjects. Bold indicates the best ACC within each train-size setting.

Dataset	Model	5%		10%		50%		70%	
		Cost	ACC	Cost	ACC	Cost	ACC	Cost	ACC
COGBCI	EEGNet	12.06 ± 2.17	0.2147 ± 0.0339	21.87 ± 2.30	0.2210 ± 0.0257	237.14 ± 0.98	0.6270 ± 0.1550	287.80 ± 0.54	0.6438 ± 0.1490
	EEGConformer	53.63 ± 8.48	0.2893 ± 0.0642	115.20 ± 7.76	0.3400 ± 0.1007	685.24 ± 3.67	0.5277 ± 0.1136	1045.37 ± 1.85	0.5954 ± 0.1043
	LaBraM	63.33 ± 13.01	0.3567 ± 0.0688	90.40 ± 22.76	0.4442 ± 0.0911	323.31 ± 84.08	0.7285 ± 0.0972	378.34 ± 119.40	0.7813 ± 0.0945
	BIOT	85.41 ± 32.46	$\mathbf{0.7041 \pm 0.0758}$	92.28 ± 37.44	$\mathbf{0.7856 \pm 0.0612}$	336.35 ± 96.97	$\mathbf{0.8932 \pm 0.0373}$	288.76 ± 102.02	$\mathbf{0.9045 \pm 0.0398}$
M3CV	EEGNet	51.50 ± 7.47	0.2208 ± 0.0344	96.98 ± 7.04	0.2469 ± 0.0312	515.89 ± 2.95	0.3492 ± 0.0776	757.33 ± 1.47	0.3938 ± 0.0866
	EEGConformer	29.29 ± 6.65	0.2442 ± 0.0147	55.52 ± 6.65	0.2414 ± 0.0196	410.75 ± 3.81	0.2859 ± 0.0705	764.47 ± 1.97	0.3259 ± 0.0967
	LaBraM	35.14 ± 12.25	0.2536 ± 0.0271	24.15 ± 8.41	0.2808 ± 0.0385	107.42 ± 41.30	0.4408 ± 0.0728	122.56 ± 32.79	0.4979 ± 0.0771
	BIOT	28.72 ± 10.54	$\mathbf{0.3028 \pm 0.0402}$	39.92 ± 16.90	$\mathbf{0.3635 \pm 0.0505}$	115.54 ± 42.79	$\mathbf{0.4791 \pm 0.0611}$	128.38 ± 56.40	$\mathbf{0.5282 \pm 0.0630}$

Table A12: **Runtime cost comparison of Rest-Tuning and Task-SFT.**

Dataset	**Task**	**LaBraM**		**BIOT**	
		Rest-tune (s)	**Task-SFT (s)**	**Rest-tune (s)**	**Task-SFT (s)**
M3CV (n=95)	P300	128.67	80.60	69.72	31.10
	Motor		51.10		40.08
	SSS		65.07		27.53
	TS		138.87		35.94
	Multitask		122.56		128.38
COGBCI (n=29)	MATB	150.37	208.18	80.90	178.94
	NBack		121.02		164.36
	Multitask		378.34		288.76

I Rest-Tuning and Task-SFT Runtime Comparison

We report the computational runtime of individual Rest-Tuning and Task-SFT for the LaBraM and BIOT models on the M3CV and COGBCI datasets. The results are summarized in Table A12.

J Data-Efficiency Analyses

Figures A2 and A3 provide task-level data-efficiency curves for Baseline, Rest-Full, and Rest-LoRA for both LaBraM and BIOT models. The results show that Rest-Tuning often reaches the same accuracy as the task-only baseline using fewer labeled training samples. This trend is observed across both COGBCI and M3CV tasks, suggesting that resting-state calibration reduces the amount of supervised task data required for subject-specific adaptation.

Table A13: Comparison with resting-state calibration methods on COGBCI. All methods use target-subject resting EEG. Source pairs denote additional source-subject rest–task data; target labels denote the labeled target-task training budget.

Method	Backbone	Source rest–task pairs	Target labels	Training time (min)	ACC (%)
BCM	EEGNet	28 subjects	0%	115.02	39.15 ± 3.62
BCM	EEGConformer	28 subjects	0%	122.56	47.15 ± 7.14
ResTL	EEGNet	28 subjects	0%	104.50	40.23 ± 3.06
ResTL	EEGConformer	28 subjects	0%	60.00	35.49 ± 0.40
Rest-Tuning	LaBraM	None	10%	4.07	44.42 ± 9.82
Rest-Tuning	BIOT	None	10%	2.93	78.56 ± 6.12
Rest-Tuning	LaBraM	None	70%	8.81	78.13 ± 9.45
Rest-Tuning	BIOT	None	70%	6.16	90.45 ± 3.98

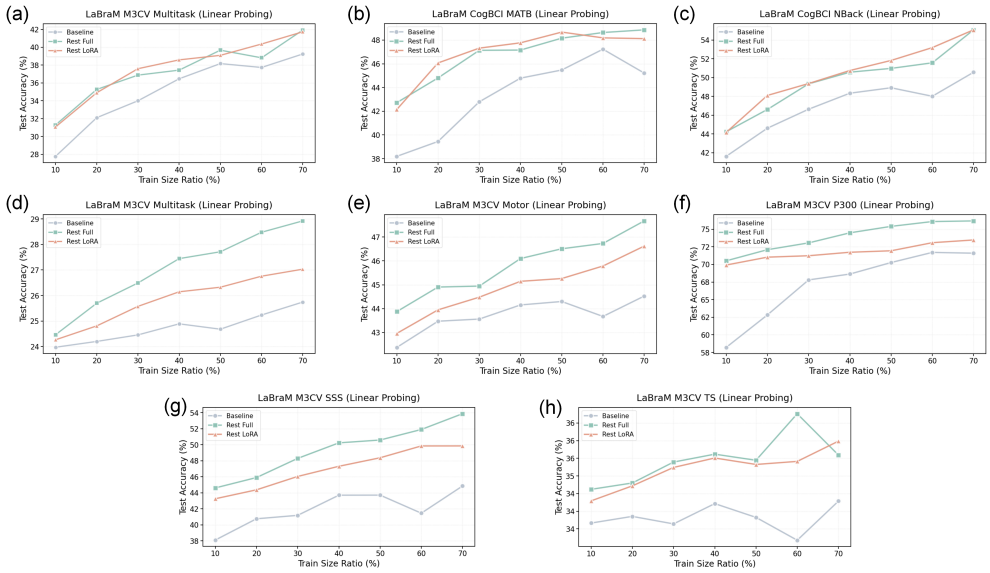

Figure A2: Data efficiency curves across downstream tasks of LaBraM model.

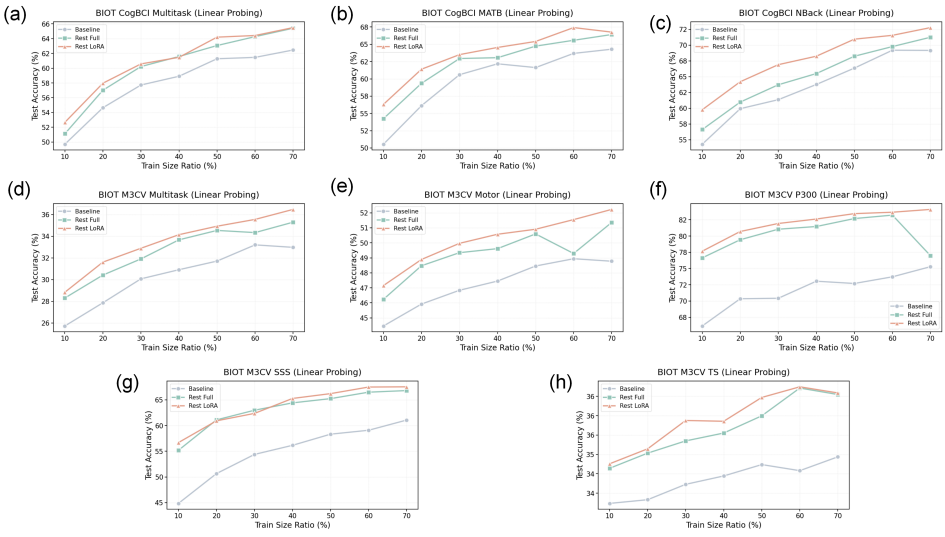

Figure A3: Data efficiency curves across downstream tasks of BIOT model.

K Additional LoRA Adapter Interpolation Analyses

In this section, we present an analysis of task-to-task interpolation, where we examine the interpolation curves between Task1→Task2 and Task2→Task1. The average performance across both directions is reflected in the combined accuracy curve shown in Figure A4. The results highlight the performance landscape of LoRA adapters in the context of multi-task adaptation.

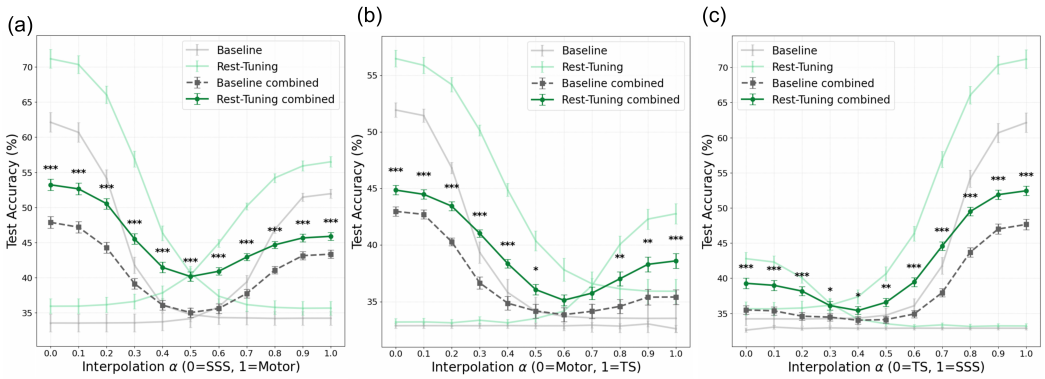

Figure A4: Detailed LoRA adapter interpolation between two tasks. Rest-Tuning (green) maintains more stable accuracy along the interpolation path compared to task-only LoRA (gray dashed).

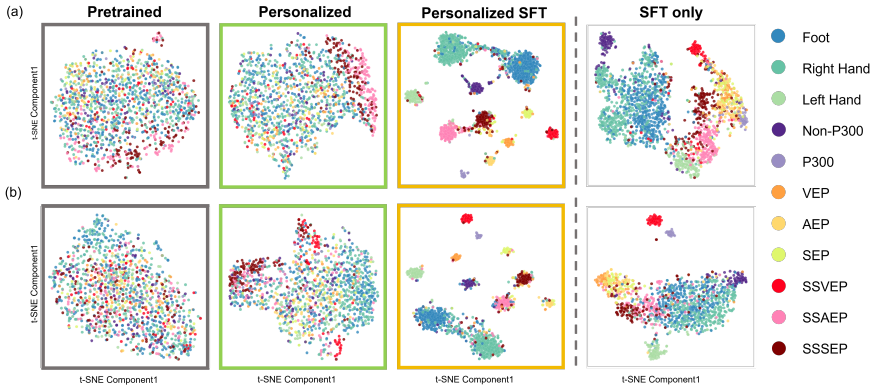

Figure A5: t-SNE visualization of LaBraM embeddings under multitask settings. We analyze two subjects using Rest-Full and SFT with 70% labeled data.

L T-SNE visualization of embeddings

We present additional t-SNE visualizations of task embeddings. Figure A5 shows multitask embedding results obtained with the LaBraM backbone on the M3CV dataset, which is analyzed for two subjects under the Rest-Full and SFT settings with 70% labeled data.

M Sensitivity to Resting-State EEG Duration

We vary the resting EEG duration from 30 to 210 seconds for LaBraM, EEGPT, and BIOT under both Rest-Full and Rest-LoRA calibration. Figure A6 presents detailed LaBraM Rest-LoRA results, while Table A14 summarizes multitask accuracy averaged over ten task-label budgets from 1% to 10%. Performance remains relatively stable across durations, with mean accuracy varying by at most 0.91 percentage points within each setting. All six configurations attain their highest mean accuracy with 60–180 seconds of rest, with no consistent benefit from longer recordings. These results support 1–3 minutes of resting EEG as a practical calibration budget.

N Detailed Performance Results

Performance under 10% and 70% training samples. We first evaluate Rest-Tuning under a low-label setting where only 10% of the task training samples are used. As shown in Figure A7, both Rest-LoRA and Rest-Full improve linear probing performance over the no-rest baseline across most EEG foundation models and downstream tasks. Table A15 further shows that these improvements persist under both uni-task and multi-task adaptation on M3CV and COGBCI. These results suggest

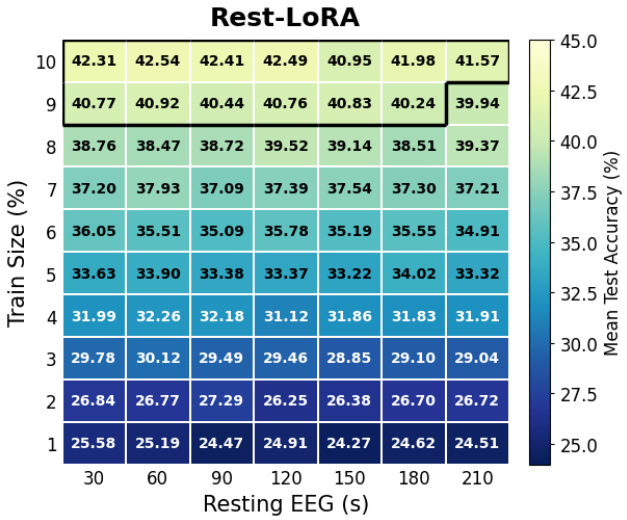

Figure A6: **Resting EEG duration vs. multitask accuracy.** LaBraM Rest-LoRA performance rapidly saturates with 1–3 minutes of resting EEG.

Table A14: Sensitivity to resting-state EEG duration across backbones. Each value is the mean accuracy (%) averaged over ten task-label budgets from 1% to 10%. Bold indicates the highest mean accuracy in each row.

Calibration	Backbone	Resting-state EEG duration						
		30 s	60 s	90 s	120 s	150 s	180 s	210 s
Rest-Full	LaBraM	34.80	35.01	35.13	**35.59**	35.11	35.22	35.42
	EEGPT	50.92	51.30	51.18	51.20	51.06	**51.37**	51.23
	BIOT	65.26	65.66	65.55	65.54	**65.80**	65.58	64.97
Rest-LoRA	LaBraM	34.29	**34.36**	34.06	34.10	33.82	33.99	33.85
	EEGPT	50.98	51.73	51.67	**51.89**	51.63	51.52	51.67
	BIOT	64.36	**64.51**	64.37	64.46	64.48	64.25	63.91

that Rest-Tuning is especially beneficial when labeled task data are limited, providing a subject-calibrated backbone that improves downstream adaptation with minimal supervision.

We also report the corresponding results under the 70% training-sample setting in Table 1. This setting provides a higher-label regime for evaluating whether the benefits of Rest-Tuning remain when more supervised task data are available. Rest-Tuning continues to improve downstream performance across different EEG foundation models, datasets, and adaptation protocols, indicating that resting-state calibration remains useful beyond the extremely low-label setting.

Detailed uni-task performance profiles. Figures A8 and A9 provide detailed uni-task performance profiles for each EEG foundation model on M3CV and COGBCI under 10% and 70% training data, respectively. The radar plots compare task-only adaptation with Rest-Tuning followed by task-specific SFT across six downstream tasks. Overall, Rest-Tuning improves the normalized task-level profiles for most backbones and adaptation settings, showing that the subject-calibrated backbone provides consistent benefits beyond the aggregate results reported in the main tables.

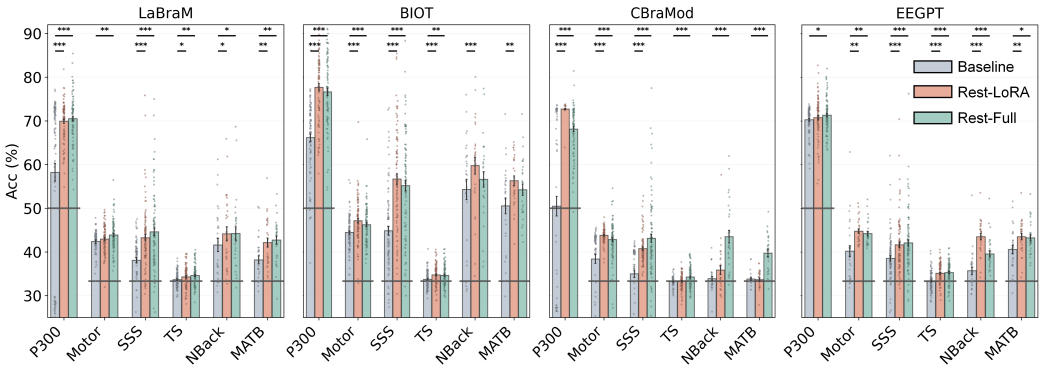

Figure A7: **Linear probing performance under 10% training data.** Bars show mean accuracy across subjects for four EEG FMs and six downstream tasks. Gray denotes the no-rest baseline, orange denotes Rest-LoRA, and green denotes Rest-Full. Error bars indicate standard deviation; significance is computed with paired subject-level tests comparing Rest-Tuning to the no-rest baseline.

Table A15: **Uni-task and multi-task** performance on M3CV and COGBCI datasets under 10% training data.

EEG Foundation Model	Rest-Tuning	Task-SFT	Trainable Params (Rest/Task)	M3CV				COGBCI			
				Unitask		Multitask		Unitask		Multitask	
				$\overline{\text{ACC}}$(%)	ΔACC(%)	$\overline{\text{ACC}}$(%)	ΔACC(%)	$\overline{\text{ACC}}$(%)	ΔACC(%)	$\overline{\text{ACC}}$(%)	ΔACC(%)
LaBraM	/	Task-LoRA	0 / 5.00%	49.36	/	25.39	/	48.66	/	37.50	/
	Rest-LoRA	Task-LoRA	5% / 5%	52.41	↑3.05	26.35	↑0.96	49.22	↑0.56	39.78	↑2.28
	Rest-Full	Task-LoRA	100% / 5%	53.13	↑3.77	26.93	↑1.54	50.06	↑1.40	40.42	↑2.92
	/	Task-Full	0 / 100%	50.81	/	26.12	/	49.48	/	39.41	/
	Rest-Full	Task-Full	100% / 100%	54.80	↑3.99	28.08	↑1.96	51.05	↑1.57	44.42	↑5.01
BIOT	/	Task-LoRA	0 / 4.5%	55.13	/	33.34	/	77.61	/	75.80	/
	Rest-LoRA	Task-LoRA	4.5% / 4.5%	61.04	↑5.91	34.19	↑0.85	79.97	↑2.36	77.69	↑1.89
	Rest-Full	Task-LoRA	100% / 4.5%	59.94	↑4.81	33.40	↑0.06	79.31	↑1.70	77.24	↑1.44
	/	Task-Full	0 / 100%	58.22	/	35.62	/	80.02	/	77.45	/
	Rest-Full	Task-Full	100% / 100%	62.29	↑4.07	36.35	↑0.73	81.37	↑1.35	78.59	↑1.14
CBraMod	/	Task-LoRA	0 / 3.03%	44.94	/	23.95	/	38.16	/	22.74	/
	Rest-LoRA	Task-LoRA	3.03% / 3.03%	47.88	↑2.94	24.50	↑0.55	46.22	↑8.06	26.00	↑3.26
	Rest-Full	Task-LoRA	100% / 3.03%	49.45	↑4.51	24.64	↑0.69	52.58	↑14.42	26.00	↑3.26
	/	Task-Full	0 / 100%	51.07	/	25.66	/	55.30	/	46.06	/
	Rest-Full	Task-Full	100% / 100%	53.12	↑2.05	26.08	↑0.42	57.96	↑2.66	50.18	↑4.12
EEGPT	/	Task-LoRA	0 / 3.03%	48.48	/	16.90	/	41.80	/	32.26	/
	Rest-LoRA	Task-LoRA	2.03% / 2.03%	49.26	↑0.78	17.78	↑0.88	42.44	↑0.64	32.54	↑0.28
	Rest-Full	Task-LoRA	100% / 2.03%	49.18	↑0.70	18.74	↑1.84	43.56	↑1.76	33.76	↑1.50
	/	Task-Full	0 / 100%	48.98	/	17.88	/	42.44	/	32.91	/
	Rest-Full	Task-Full	100% / 100%	49.56	↑0.58	19.14	↑1.26	43.33	↑0.89	34.35	↑1.44

Table A16: **Uni-task and multi-task** performance on M3CV and COGBCI datasets.

EEG foundation model	Rest-Tuning	Task-SFT	Unitask								Multitask							
			M3CV				COG-BCI				M3CV				COG-BCI			
			train 10%		train 70%		train 10%		train 70%		train 10%		train 70%		train 10%		train 70%	
			$\overline{\text{F1}}$ (%)	Δ F1(%)	$\overline{\text{F1}}$ (%)	Δ F1(%)	$\overline{\text{F1}}$ (%)	Δ F1(%)	$\overline{\text{F1}}$ (%)	Δ F1(%)	F1 (%)	Δ F1(%)	F1 (%)	Δ F1(%)	F1 (%)	Δ F1(%)	F1 (%)	Δ F1(%)
LaBraM	/	LoRA	44.92	/	55.24	/	47.92	/	59.86	/	19.68	/	34.17	/	36.84	/	56.39	/
	Rest-LoRA	LoRA	51.10	↑6.18	63.99	↑8.75	48.39	↑0.47	63.14	↑3.28	21.46	↑1.78	40.33	↑6.16	39.18	↑2.34	60.92	↑4.53
	Rest-Full	LoRA	52.00	↑7.08	64.35	↑9.11	49.48	↑1.56	68.70	↑8.84	22.22	↑2.54	41.23	↑7.06	39.68	↑2.84	66.53	↑10.14
	/	Full	48.47	/	59.97	/	48.40	/	72.19	/	21.29	/	44.23	/	38.68	/	73.24	/
	Rest-Full	Full	53.59	↑5.12	67.90	↑7.93	50.42	↑2.02	75.30	↑3.11	23.69	↑2.40	48.44	↑4.21	43.62	↑4.94	77.95	↑4.71
BIOT	/	LoRA	53.04	/	64.87	/	77.35	/	90.08	/	30.45	/	48.08	/	75.60	/	88.93	/
	Rest-LoRA	LoRA	60.62	↑7.58	70.35	↑5.48	79.86	↑2.51	90.92	↑0.84	31.73	↑1.28	50.87	↑2.79	77.53	↑1.93	89.46	↑0.53
	Rest-Full	LoRA	59.61	↑6.57	67.63	↑2.76	79.17	↑1.82	91.19	↑1.11	31.25	↑0.80	49.21	↑1.13	77.06	↑1.46	89.58	↑0.65
	/	Full	56.23	/	66.82	/	79.92	/	90.38	/	32.97	/	49.32	/	77.24	/	89.33	/
	Rest-Full	Full	61.93	↑5.70	68.85	↑2.03	81.25	↑1.33	91.23	↑0.85	34.40	↑1.43	51.45	↑2.13	78.38	↑1.14	90.44	↑1.11
CBraMod	/	LoRA	37.25	/	40.98	/	34.81	/	45.76	/	12.40	/	14.51	/	19.36	/	36.98	/
	Rest-LoRA	LoRA	41.36	↑4.11	46.96	↑5.98	44.39	↑9.58	59.62	↑13.86	14.36	↑1.96	16.17	↑1.66	25.54	↑6.18	39.71	↑2.73
	Rest-Full	LoRA	45.38	↑8.13	54.04	↑13.06	51.81	↑17.00	59.94	↑14.18	15.58	↑3.18	20.42	↑5.91	25.54	↑6.18	47.93	↑10.95
	/	Full	43.99	/	56.93	/	52.89	/	70.71	/	15.36	/	30.60	/	42.21	/	70.23	/
	Rest-Full	Full	52.02	↑8.03	62.05	↑5.12	57.36	↑4.47	75.24	↑4.53	20.20	↑4.84	35.67	↑5.07	49.16	↑6.95	73.44	↑3.21
EEGPT	/	LoRA	44.58	/	49.38	/	41.39	/	54.39	/	13.89	/	25.99	/	32.14	/	48.28	/
	Rest-LoRA	LoRA	47.53	↑2.95	52.79	↑3.41	42.30	↑0.91	54.70	↑0.31	15.16	↑1.27	26.19	↑0.20	32.28	↑0.14	48.58	↑0.30
	Rest-Full	LoRA	47.49	↑2.91	54.27	↑4.89	43.49	↑2.10	54.63	↑0.24	16.09	↑2.20	26.88	↑0.89	33.41	↑1.27	49.67	↑1.39
	/	Full	45.42	/	52.43	/	41.78	/	54.25	/	15.65	/	25.30	/	32.80	/	47.60	/
	Rest-Full	Full	48.08	↑2.66	54.78	↑2.35	42.90	↑1.12	57.06	↑2.81	16.92	↑1.27	28.69	↑3.39	34.01	↑1.21	57.67	↑10.07

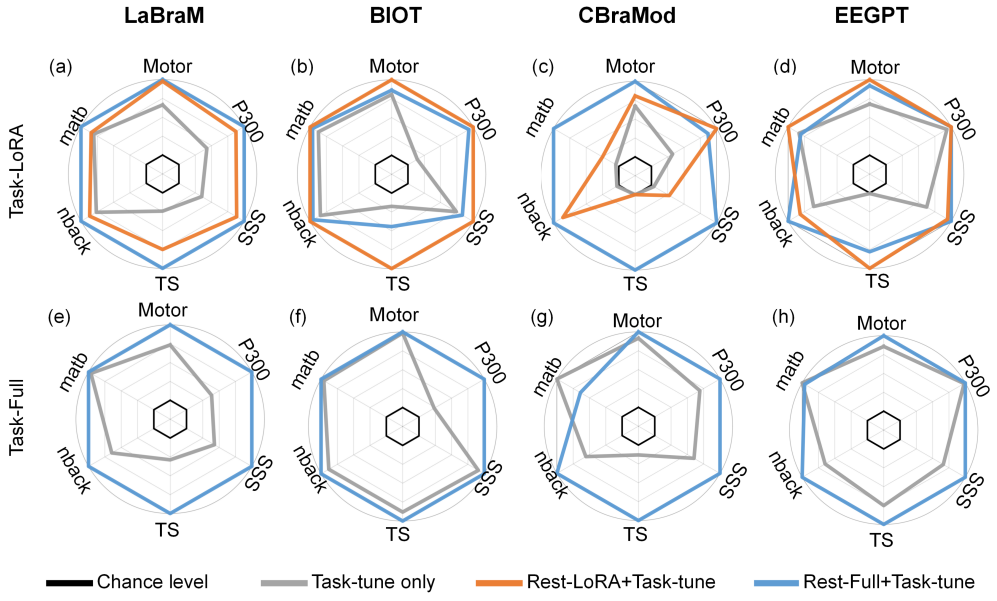

Figure A8: **Uni-task performance profiles on M3CV under 10% training data.** We display normalized accuracy across six tasks using Rest-Tuning on four different EEG FM backbones. Task-specific SFT adaptation strategies, including LoRA and Full tuning, are used after Rest-Tuning.

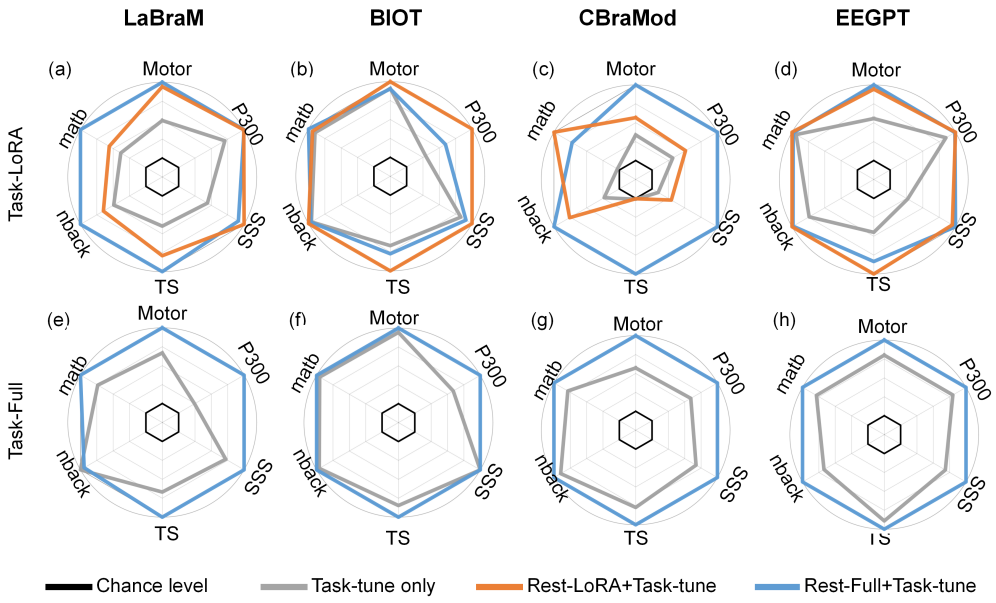

Figure A9: **Uni-task performance profiles on M3CV under 70% training data.** We display normalized accuracy across six tasks using Rest-Tuning on four different EEG FM backbones.

Table A17: Uni-task performance on M3CV and COGBCI datasets (training size = 10%)

EEG foundation model	Rest-Tuning	Task-SFT	**Motor** (3-Class)		**P300** (2-Class)		**Steady State Sensory** (3-Class)		**Transient Sensory** (3-Class)		**Nback** (3-Class)		**MATB** (3-Class)	
			ACC	**F1**	**ACC**	**F1**	**ACC**	**F1**	**ACC**	**F1**	**ACC**	**F1**	**ACC**	**F1**
LaBraM	/	Task-LP	0.4238±0.0380	0.3348±0.0657	0.5821±0.1936	0.4884±0.2273	0.3811±0.0570	0.2773±0.0786	0.3367±0.0172	0.2724±0.0629	0.4161±0.0824	0.3867±0.0985	0.3819±0.0493	0.3326±0.0962
	Rest-LoRA	Task-LP	0.4296±0.0297	0.3994±0.0341	0.6991±0.0469	0.6349±0.0454	0.4317±0.0763	0.3883±0.0835	0.3429±0.0239	0.3226±0.0329	0.4415±0.0869	0.4335±0.0884	0.4214±0.0554	0.4109±0.0600
	Rest-Full	Task-LP	0.4388±0.0339	0.4070±0.0384	0.6497±0.1121	0.6123±0.1099	0.4461±0.0883	0.4043±0.0940	0.3462±0.0234	0.3291±0.0341	0.4422±0.0813	0.4263±0.0835	0.4272±0.0506	0.4220±0.0530
	/	Task-LoRA	0.4473±0.0268	0.4076±0.0545	0.7099±0.0391	0.6302±0.0361	0.4696±0.0938	0.4404±0.1080	0.3474±0.0236	0.3184±0.0520	0.5089±0.1088	0.5038±0.1074	0.4642±0.0542	0.4546±0.0622
	Rest-LoRA	Task-LoRA	0.4629±0.0368	0.4485±0.0437	0.7663±0.0656	0.7510±0.0759	0.5132±0.1082	0.5007±0.1132	0.3538±0.0242	0.3438±0.0284	0.5163±0.1049	0.5099±0.1020	0.4681±0.0421	0.4578±0.0423
	Rest-Full	Task-LoRA	0.4640±0.0386	0.4508±0.0425	0.7785±0.0644	0.7651±0.0741	0.5268±0.1032	0.5153±0.1075	0.3559±0.0262	0.3486±0.0269	0.5253±0.1006	0.5183±0.0999	0.4759±0.0616	0.4712±0.0619
	/	Task-Full	0.4618±0.0371	0.4235±0.0617	0.7251±0.0293	0.7333±0.0870	0.4859±0.0988	0.4539±0.1196	0.3596±0.0271	0.3278±0.0490	0.5212±0.1097	0.5134±0.1135	0.4684±0.0536	0.4546±0.0593
	Rest-Full	Task-Full	0.4758±0.0412	0.4620±0.0458	0.8096±0.0796	0.7959±0.0901	0.5358±0.1192	0.5232±0.1239	0.3708±0.0289	0.3626±0.0311	0.5509±0.1157	0.5461±0.1165	0.4701±0.0654	0.4623±0.0657
BIOT	/	Task-LP	0.4446±0.0375	0.4154±0.0541	0.6616±0.0818	0.6240±0.0684	0.4485±0.0965	0.4084±0.1168	0.3373±0.0199	0.3081±0.0442	0.5430±0.1249	0.5243±0.1309	0.5055±0.0967	0.4810±0.1211
	Rest-LoRA	Task-LP	0.4715±0.0507	0.4619±0.0532	0.7767±0.0820	0.7612±0.0901	0.5669±0.1140	0.5573±0.1183	0.3476±0.0230	0.3412±0.0248	0.5976±0.1045	0.5896±0.1044	0.5631±0.0582	0.5553±0.0592
	Rest-Full	Task-LP	0.4624±0.0437	0.4543±0.0471	0.7663±0.0808	0.7568±0.0852	0.5518±0.1179	0.5438±0.1219	0.3464±0.0235	0.3403±0.0257	0.5663±0.0941	0.5550±0.0990	0.5425±0.0697	0.5354±0.0707
	/	Task-LoRA	0.5063±0.0478	0.4975±0.0556	0.6965±0.0631	0.6518±0.0697	0.6561±0.1246	0.6497±0.1303	0.3464±0.0259	0.3225±0.0508	0.7947±0.1031	0.7912±0.1055	0.7574±0.0606	0.7558±0.0615
	Rest-LoRA	Task-LoRA	0.5192±0.0459	0.5153±0.0484	0.8757±0.0653	0.8726±0.0691	0.6900±0.1083	0.6868±0.1122	0.3566±0.0242	0.3501±0.0250	0.8220±0.0936	0.8217±0.0935	0.7773±0.0686	0.7754±0.0691
	Rest-Full	Task-LoRA	0.5103±0.0497	0.5063±0.0525	0.8672±0.0746	0.8650±0.0755	0.6690±0.1111	0.6656±0.1137	0.3511±0.0238	0.3474±0.0242	0.8146±0.0944	0.8138±0.0950	0.7715±0.0745	0.7696±0.0743
	/	Task-Full	0.5308±0.0488	0.5229±0.0564	0.7382±0.0706	0.7002±0.0872	0.7031±0.1247	0.6951±0.1356	0.3567±0.0242	0.3311±0.0450	0.8104±0.0876	0.8107±0.0881	0.7899±0.0569	0.7877±0.0582
	Rest-Full	Task-Full	0.5319±0.0465	0.5299±0.0479	0.8871±0.0635	0.8835±0.0681	0.7150±0.1130	0.7124±0.1147	0.3577±0.0224	0.3513±0.0258	0.8309±0.0936	0.8301±0.0940	0.7965±0.0664	0.7949±0.0666
CBraMod	/	Task-LP	0.3839±0.0980	0.2543±0.0820	0.5048±0.2180	0.3857±0.24105	0.3501±0.0796	0.2155±0.0785	0.3325±0.0098	0.1990±0.0439	0.3389±0.0271	0.2085±0.0520	0.3379±0.0154	0.2092±0.0522
	Rest-LoRA	Task-LP	0.4381±0.0157	0.3856±0.0362	0.7269±0.0094	0.6141±0.0063	0.4076±0.0348	0.3384±0.0440	0.3328±0.0170	0.2897±0.0298	0.3586±0.0551	0.3011±0.0656	0.3392±0.0221	0.2991±0.0423
	Rest-Full	Task-LP	0.4283±0.0407	0.3968±0.0419	0.6811±0.0550	0.6257±0.0450	0.4314±0.0893	0.3885±0.0954	0.3428±0.0240	0.3146±0.0349	0.4675±0.0947	0.4374±0.1018	0.4009±0.0469	0.3749±0.0532
	/	Task-LoRA	0.4334±0.0134	0.3406±0.0523	0.6556±0.0970	0.5960±0.0719	0.3749±0.0453	0.2935±0.0606	0.3335±0.0128	0.2600±0.0468	0.3994±0.0730	0.3651±0.0915	0.3637±0.0333	0.3310±0.0430
	Rest-LoRA	Task-LoRA	0.4397±0.0291	0.3902±0.0456	0.7264±0.0162	0.6140±0.0091	0.4140±0.0485	0.3500±0.0554	0.3351±0.0151	0.3000±0.0341	0.5377±0.1565	0.5220±0.1667	0.3866±0.0685	0.3657±0.0675
	Rest-Full	Task-LoRA	0.4475±0.0406	0.4135±0.0472	0.7159±0.0652	0.6476±0.0751	0.4575±0.1189	0.4140±0.1222	0.3570±0.0394	0.3401±0.0477	0.5730±0.1430	0.5653±0.1494	0.4785±0.0937	0.4708±0.0957
	/	Task-Full	0.4597±0.0370	0.3962±0.0631	0.7366±0.0376	0.6365±0.0643	0.4956±0.0974	0.4332±0.1188	0.3510±0.0244	0.2937±0.0496	0.5654±0.1206	0.5454±0.1253	0.5406±0.0924	0.5123±0.1029
	Rest-Full	Task-Full	0.4635±0.0415	0.4500±0.0466	0.7677±0.0722	0.7566±0.0797	0.5256±0.1021	0.5149±0.1071	0.3680±0.0236	0.3593±0.0254	0.6329±0.1195	0.6279±0.1222	0.5262±0.0902	0.5192±0.0910
EEGPT	/	Task-LP	0.4019±0.0671	0.3899±0.0813	0.7028±0.0304	0.6187±0.0170	0.3854±0.0597	0.3335±0.0734	0.3338±0.0216	0.2818±0.0561	0.3572±0.0445	0.3392±0.0623	0.4058±0.0484	0.3951±0.0670
	Rest-LoRA	Task-LP	0.4445±0.0291	0.4168±0.0338	0.7075±0.0501	0.6467±0.0494	0.4165±0.0623	0.3865±0.0652	0.3514±0.0216	0.3482±0.0223	0.3820±0.0376	0.3782±0.0403	0.4349±0.0389	0.4333±0.0391
	Rest-Full	Task-LP	0.4408±0.0264	0.4156±0.0300	0.7133±0.0369	0.6563±0.0462	0.4210±0.0673	0.3954±0.0683	0.3528±0.0229	0.3496±0.0250	0.3957±0.0392	0.3913±0.0416	0.4325±0.0379	0.4296±0.0355
	/	Task-LoRA	0.4560±0.0693	0.4530±0.0733	0.7219±0.0376	0.6466±0.0548	0.4226±0.0732	0.3803±0.0860	0.3383±0.0259	0.3032±0.0452	0.3868±0.0409	0.3791±0.0445	0.4492±0.0532	0.4486±0.0537
	Rest-LoRA	Task-LoRA	0.4492±0.0315	0.4306±0.0359	0.7270±0.0445	0.6994±0.0555	0.4359±0.0699	0.4173±0.0760	0.3582±0.0247	0.3538±0.0261	0.3922±0.0321	0.3903±0.0330	0.4565±0.0376	0.4557±0.0372
	Rest-Full	Task-LoRA	0.4461±0.0333	0.4299±0.0377	0.7270±0.0445	0.6974±0.0556	0.4378±0.0755	0.4201±0.0813	0.3562±0.0245	0.3521±0.0257	0.4080±0.0382	0.4065±0.0390	0.4632±0.0341	0.4633±0.0337
	/	Task-Full	0.4449±0.0272	0.4017±0.0515	0.7297±0.0501	0.6771±0.0699	0.4280±0.0658	0.3980±0.0775	0.3564±0.0274	0.3400±0.0392	0.3979±0.0370	0.3873±0.0387	0.4508±0.0413	0.4483±0.0418
	Rest-Full	Task-Full	0.4508±0.0328	0.4369±0.0370	0.7311±0.0504	0.7056±0.0571	0.4419±0.0696	0.4299±0.0731	0.3587±0.0244	0.3508±0.0280	0.4059±0.0383	0.4029±0.0389	0.4607±0.0369	0.4550±0.0399

Table A18: Uni-task performance on M3CV and COGBCI datasets (training size = 70%)

EEG foundation model	Rest-Tuning	Task-SFT	**Motor** (3-Class)		**P300** (2-Class)		**Steady State Sensory** (3-Class)		**Transient Sensory** (3-Class)		**Nback** (3-Class)		**MATB** (3-Class)	
			ACC	**F1**	**ACC**	**F1**	**ACC**	**F1**	**ACC**	**F1**	**ACC**	**F1**	**ACC**	**F1**
LaBraM	/	Task-LP	0.4453±0.0456	0.3913±0.0651	0.7160±0.0219	0.6211±0.0274	0.4485±0.1004	0.3840±0.1124	0.3429±0.0390	0.3165±0.0621	0.5056±0.1143	0.4976±0.1151	0.4522±0.0534	0.4891±0.0556
	Rest-LoRA	Task-LP	0.4662±0.0443	0.4389±0.0514	0.7347±0.0621	0.6647±0.0824	0.4908±0.1018	0.4588±0.1075	0.3599±0.0418	0.3487±0.0440	0.5508±0.1058	0.5456±0.1054	0.4812±0.0542	0.4715±0.0528
	Rest-Full	Task-LP	0.4767±0.0546	0.4556±0.0636	0.7618±0.0748	0.7123±0.0976	0.5330±0.1146	0.5073±0.1234	0.3559±0.0434	0.3439±0.0471	0.5507±0.0960	0.5410±0.0995	0.4886±0.0561	0.4803±0.0574
	/	Task-LoRA	0.5198±0.0616	0.5111±0.0664	0.7593±0.0597	0.7151±0.0868	0.6210±0.1359	0.6156±0.1385	0.3783±0.0458	0.3676±0.0506	0.6479±0.1091	0.6428±0.1108	0.5613±0.0745	0.5543±0.0640
	Rest-LoRA	Task-LoRA	0.5644±0.0728	0.5617±0.0737	0.9078±0.0620	0.9059±0.0646	0.7117±0.1298	0.7078±0.1321	0.3898±0.0438	0.3843±0.0459	0.6786±0.1127	0.6746±0.1142	0.5919±0.0741	0.5881±0.0777
	Rest-Full	Task-LoRA	0.5690±0.0714	0.5654±0.0745	0.9084±0.0583	0.9060±0.0618	0.7176±0.1264	0.7148±0.1286	0.3944±0.0486	0.3878±0.0499	0.7319±0.1166	0.7294±0.1172	0.6469±0.0863	0.6445±0.0879
	/	Task-Full	0.5765±0.0691	0.5662±0.0785	0.7792±0.0713	0.7457±0.0988	0.7100±0.1412	0.7040±0.1464	0.4008±0.0433	0.3828±0.0550	0.7636±0.1232	0.7574±0.1277	0.6879±0.0835	0.6864±0.0834
	Rest-Full	Task-Full	0.6114±0.0741	0.6074±0.0768	0.9472±0.0508	0.9457±0.0545	0.7537±0.1301	0.7482±0.1356	0.4222±0.0461	0.4146±0.0492	0.7844±0.1281	0.7823±0.1291	0.7260±0.1123	0.7236±0.1141
BIOT	/	Task-LP	0.4829±0.0546	0.4680±0.0611	0.7528±0.1185	0.7318±0.1296	0.6107±0.1145	0.6007±0.1232	0.3428±0.0402	0.3241±0.0568	0.6583±0.1080	0.6503±0.1111	0.6234±0.0686	0.6174±0.0711
	Rest-LoRA	Task-LP	0.5223±0.0562	0.5166±0.0590	0.8409±0.0675	0.8288±0.0804	0.6754±0.1084	0.6704±0.1109	0.3659±0.0410	0.3594±0.0423	0.7277±0.0926	0.7239±0.0939	0.6676±0.0621	0.6623±0.0661
	Rest-Full	Task-LP	0.5136±0.0621	0.5096±0.0642	0.7698±0.1319	0.7578±0.1335	0.6800±0.1116	0.6630±0.1141	0.3655±0.0411	0.3611±0.0418	0.7124±0.1046	0.7078±0.1058	0.6641±0.0671	0.6587±0.0709
	/	Task-LoRA	0.6022±0.0630	0.5973±0.0657	0.7947±0.0867	0.7812±0.0976	0.8463±0.0870	0.8454±0.0879	0.3793±0.0406	0.3710±0.0421	0.9148±0.0588	0.9145±0.0591	0.8880±0.0459	0.8871±0.0466
	Rest-LoRA	Task-LoRA	0.6121±0.0631	0.6109±0.0640	0.9451±0.0427	0.9441±0.0440	0.8795±0.0806	0.8786±0.0816	0.3860±0.0491	0.3803±0.0498	0.9238±0.0457	0.9234±0.0462	0.8958±0.0526	0.8949±0.0541
	Rest-Full	Task-LoRA	0.6029±0.0662	0.6015±0.0676	0.8740±0.0940	0.8667±0.1026	0.8620±0.0885	0.8613±0.0892	0.3817±0.0392	0.3755±0.0401	0.9165±0.0574	0.9163±0.0578	0.9077±0.0453	0.9075±0.0455
	/	Task-Full	0.6158±0.0679	0.6132±0.0686	0.8306±0.0802	0.8239±0.0867	0.8712±0.0841	0.8698±0.0859	0.3807±0.0423	0.3660±0.0491	0.9128±0.0567	0.9130±0.0565	0.8950±0.0526	0.8945±0.0526
	Rest-Full	Task-Full	0.6214±0.0629	0.6198±0.0638	0.8947±0.0879	0.8874±0.0979	0.8713±0.0884	0.8706±0.0892	0.3835±0.0355	0.3763±0.0375	0.9208±0.0516	0.9206±0.0519	0.9040±0.0421	0.9039±0.0419
CBraMod	/	Task-LP	0.4200±0.0537	0.2838±0.0550	0.6097±0.1756	0.5153±0.1852	0.3881±0.0610	0.2567±0.0744	0.3367±0.0180	0.2097±0.0501	0.3742±0.0376	0.3028±0.0653	0.3386±0.0213	0.2216±0.0537
	Rest-LoRA	Task-LP	0.4384±0.0265	0.3836±0.0434	0.7279±0.0021	0.6132±0.0028	0.4138±0.0614	0.3491±0.0629	0.3330±0.0275	0.2642±0.0416	0.3827±0.0606	0.3194±0.0619	0.3405±0.0256	0.2840±0.0370
	Rest-Full	Task-LP	0.4557±0.0472	0.4258±0.0524	0.7070±0.0558	0.6414±0.0614	0.4816±0.0845	0.4425±0.0928	0.3589±0.0390	0.3444±0.0418	0.5267±0.0877	0.5069±0.0897	0.4479±0.0669	0.4281±0.0675
	/	Task-LoRA	0.4376±0.0227	0.3586±0.0528	0.7089±0.0375	0.6134±0.0181	0.4298±0.0581	0.3539±0.0699	0.3444±0.0357	0.3132±0.0496	0.5511±0.1511	0.5369±0.1626	0.4041±0.0936	0.3783±0.1003
	Rest-LoRA	Task-LoRA	0.4571±0.0699	0.4297±0.0716	0.7451±0.0969	0.6954±0.1190	0.4745±0.1061	0.4355±0.1206	0.3417±0.0377	0.3176±0.0443	0.6567±0.1295	0.6452±0.1316	0.5622±0.1156	0.5472±0.1184
	Rest-Full	Task-LoRA	0.4829±0.0661	0.4690±0.0690	0.8064±0.0781	0.7964±0.0863	0.5454±0.1206	0.5237±0.1284	0.3833±0.0561	0.3726±0.0561	0.6754±0.1252	0.6675±0.1258	0.5447±0.0998	0.5313±0.1012
	/	Task-Full	0.4957±0.0569	0.4641±0.0749	0.8335±0.0655	0.8257±0.0756	0.6295±0.1175	0.6114±0.1320	0.3967±0.0474	0.3758±0.0620	0.7686±0.0994	0.7630±0.1026	0.6674±0.0834	0.6512±0.0960
	Rest-Full	Task-Full	0.5288±0.0662	0.5236±0.0673	0.8970±0.0659	0.8966±0.0655	0.6713±0.1241	0.6660±0.1286	0.4024±0.0489	0.3957±0.0513	0.8046±0.0933	0.8035±0.0938	0.7033±0.0796	0.7012±0.0802
EEGPT	/	Task-LP	0.4417±0.0428	0.4004±0.0527	0.7226±0.0361	0.6444±0.0537	0.4168±0.0763	0.3676±0.0793	0.3498±0.0410	0.3323±0.0534	0.4793±0.0669	0.4760±0.0674	0.5054±0.0487	0.5017±0.0484
	Rest-LoRA	Task-LP	0.4712±0.0561	0.4545±0.0624	0.7656±0.0590	0.7280±0.0804	0.4904±0.0942	0.4740±0.1023	0.3774±0.0443	0.3731±0.0454	0.5412±0.0565	0.5362±0.0597	0.5432±0.0604	0.5413±0.0607
	Rest-Full	Task-LP	0.4670±0.0561	0.4503±0.0628	0.7727±0.0679	0.7411±0.0829	0.4978±0.0988	0.4852±0.1037	0.3778±0.0415	0.3744±0.0431	0.5282±0.0547	0.5233±0.0560	0.5235±0.0599	0.5179±0.0624
	/	Task-LoRA	0.4596±0.0572	0.4283±0.0678	0.7895±0.0650	0.7681±0.0842	0.4613±0.1100	0.4220±0.1236	0.3707±0.0444	0.3566±0.0578	0.5392±0.0834	0.5360±0.0828	0.5562±0.0590	0.5518±0.0619
	Rest-LoRA	Task-LoRA	0.5544±0.1094	0.5402±0.1158	0.7882±0.0791	0.7710±0.0896	0.4328±0.1175	0.4051±0.1211	0.4063±0.0818	0.3953±0.0875	0.5541±0.0887	0.5528±0.0887	0.5443±0.0607	0.5411±0.0616
	Rest-Full	Task-LoRA	0.4870±0.0553	0.4803±0.0582	0.7971±0.0718	0.7944±0.0779	0.5319±0.1061	0.5219±0.1221	0.3793±0.0410	0.3743±0.0426	0.5423±0.0822	0.5395±0.0841	0.5574±0.0579	0.5531±0.0590
	/	Task-Full	0.4783±0.0495	0.4536±0.0612	0.7942±0.0663	0.7804±0.0766	0.5125±0.0980	0.4972±0.1057	0.3804±0.0485	0.3658±0.0571	0.5442±0.0849	0.5373±0.0859	0.5568±0.0653	0.5477±0.0736
	Rest-Full	Task-Full	0.4894±0.0577	0.4812±0.0613	0.8171±0.0686	0.8063±0.0755	0.5367±0.1035	0.5265±0.1085	0.3822±0.0424	0.3773±0.0427	0.5741±0.0735	0.5706±0.0741	0.5753±0.0737	0.5706±0.0755

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introduction state the main contribution of Rest-Tuning as a resting-state calibration strategy for personalizing EEG foundation models. The claims are supported by experiments across multiple EEG datasets and tasks, comparisons with task-tuning baselines, and representation-level mechanism analyses (Section 4).

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: The Discussion and Limitations section explicitly addresses that the temporal evaluation is limited to two weeks under the COGBCI acquisition protocol. Longer-term stability and robustness to changes in neural state, electrode contact, and recording conditions remain to be investigated. (Section 5).

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [N/A].

   Justification: No formal theoretical results, theorems, or proofs are presented; equations are for defining objectives and evaluation metrics.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: Justification: Detailed dataset descriptions, preprocessing, model architectures, Rest-Tuning and task adaptation procedures, hyperparameters, metrics, and statistical tests are provided (Sections 3- 4, Appendix C).

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [Yes]

   Justification: We will release code, scripts, and instructions; datasets are accessed through their original public providers.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [Yes]

   Justification: The paper details training/test splits, preprocessing, model hyperparameters, LoRA settings, mask ratios, and adaptation protocol (Sections 4.1 and Table A4).

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

   Justification: Figure 2, Figure 3, Figure 4 and Figure A7include standard deviation error bars and paired t-tests for significance.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: Runtime and task-adaptation cost are reported (Figure 7, Table 2, Table A3 and Table A13), showing per-subject Rest-Tuning time of 1–3 minutes and total adaptation cost for different backbones.

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: Only public EEG datasets are used, data is anonymized, and no ethics violations occur.

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

    Justification: Potential positive impacts (improved BCI personalization, reduced calibration burden) and negative impacts (misuse, error-prone outputs) and mitigations are discussed (Discussion, Section 5).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [Yes]

    Justification: Data is public, Rest-Tuning introduces low-risk models, and calibration is reusable across tasks.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

    Justification: EEG datasets and EEG foundation models are cited with original papers and public licenses.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [N/A]

    Justification:No new datasets are released; only code will be provided.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: The study uses existing EEG datasets; no new participants or crowdsourcing are involved.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

Answer: [N/A]

Justification: Only existing public EEG datasets are used; no new human subjects research.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

Answer: [N/A]

Justification: No LLMs are used in the method; only deep learning models for EEG processing.